Zhigeng Pan Xiaopeng Zhang
Abdennour El Rhalibi Woontack Woo
Yi Li (Eds.)

LNCS 5093

Technologies for E-Learning and Digital Entertainment

Third International Conference, Edutainment 2008
Nanjing, China, June 2008
Proceedings

Springer

Lecture Notes in Computer Science 5093

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Zhigeng Pan Xiaopeng Zhang
Abdennour El Rhalibi Woontack Woo
Yi Li (Eds.)

Technologies for E-Learning and Digital Entertainment

Third International Conference, Edutainment 2008
Nanjing, China, June 25-27, 2008
Proceedings

Springer

Volume Editors

Zhigeng Pan
Zhejiang University, State Key Lab of CAD&CG
Hangzhou 310027,China
E-mail: zgpan@cad.zju.edu.cn

Xiaopeng Zhang
Chinese Academy of Sciences
Institute of Automation, National Laboratory of Pattern Recognition
Beijing 100080, China
E-mail: xpzhang@nlpr.ia.ac.cn

Abdennour El Rhalibi
Liverpool John Moores University, School of Computing and Mathematical Sciences
Liverpool L3 3AF, UK
E-mail: A.Elrhalibi@ljmu.ac.uk

Woontack Woo
Gwangju Institute of Science and Technology (GIST), U-VR Lab
Gwangju 500-712, South Korea
E-mail: wwoo@gist.ac.kr

Yi Li
Nanjing Normal University, Edu-game Research Center, Nanjing, China
E-mail: yilisd@163.com

Library of Congress Control Number: 2008929344

CR Subject Classification (1998): K.3.1-2, I.2.1, H.5, H.3, I.3

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

ISSN 0302-9743
ISBN-10 3-540-69734-9 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-69734-3 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12279272 06/3180 5 4 3 2 1 0

Preface

With the widespread interest in digital entertainment and the advances in the technologies of computer graphics, multimedia and virtual reality technologies, a new area—"Edutainment"—has been accepted as a union of education and computer entertainment. Edutainment is recognized as an effective way of learning through a medium, such as a computer, software, games or VR applications, that both educates and entertains.

The Edutainment conference series was established and followed as a special event for the new interests in e-learning and digital entertainment. The main purpose of Edutainment conferences is the discussion, presentation, and information exchange of scientific and technological developments in the new community. The Edutainment conference series is a very interesting opportunity for researchers, engineers and graduate students who wish to communicate at these international annual events. The conference series includes plenary invited talks, workshops, tutorials, paper presentation tracks and panel discussions. The Edutainment conference series was initiated in Hangzhou, China in 2006. Following the success of the first event (Edutainment 2006 in Hangzhou, China) and the second one (Edutainment 2007 in Hong Kong, China), Edutainment 2008 was held June 25–27, 2007 in Nanjing, China.

This year, we received 219 submissions from 26 different countries and regions, including United Arab Emirates, Canada, Thailand, New Zealand, Austria, Turkey, Germany, Switzerland, Brazil, Cuba, Australia, Hong Kong (China), Pakistan, Mexico, Czech Republic, USA, Malaysia, Italy, Spain, France, UK, The Netherlands, Taiwan (China), Japan, South Korea, and China. A total of 83 papers were selected, after peer review, for this volume. Topics of these papers fall into ten different areas ranging from fundamental issues in geometric modeling and imaging to virtual reality systems and their applications in computer entertainment and education. These topics include E-Learning Platforms and Tools, E-Learning System for Education, Application of E-Learning Systems, E-Learning Resource Management, Interaction in Game and Education, Integration of Game and Education, Game Design and Development, Virtual Characters, Animation and Navigation, Graphics Rendering and Digital Media, and Geometric Modeling for Games and Virtual Reality.

We are grateful to the International Program Committee and the reviewers for their great effort and serious work to get all the papers reviewed in a short period of time. We are grateful to the Organizing Committee and Executive Committee for their support of this event. We would also like to thank the authors and participants for their enthusiasm and contribution to the success of this conference.

The success of Edutainment 2008 was also due to the financial and practical support of various institutions.

Sponsors

- VR Committee, China Society of Image and Graphics
- Zhejiang University

Co-sponsors

- National Science Foundation of China (NSFC)
- *International Journal of Virtual Reality* (IJVR)
- *International Journal of Computer Games Technology* (IJCGT)
- *Transactions on Edutainment* (ToE)
- Nanjing Normal University, China
- Hohai University, China
- LIAMA-NLPR, Institute of Automation, CAS, China

We would like to thank all of them for offering the opportunity to organize Edutainment 2008 in a way that provided a diversified scientific and social program. Especially, we would like to thank all members of the International Program Committee and Organizing Committee for their great job in defining the conference topics, reviewing the large number of submitted papers, and managing to put all the material together for this great event.

March 2008

Zhigeng Pan
Xiaopeng Zhang
Abdennour El Rhalibi
Woontack Woo
Yi Li

Organization

Committee

Conference Honorary Chairs

Ruth Aylett, Heriot-Watt University, UK
Newton Lee, ACM Computers in Entertainment, USA
Yongzhong Song, Nanjing Normal University, China

Conference Co-chairs

Zhigeng Pan, Zhejiang University, China
Jim Chen, George Mason University, USA
Ryohei Nakatsu, Kwansei Gakuin University, Japan

Program Co-chairs

Xiaopeng Zhang, Institute of Automation, CAS, China
Abdennour El Rhalibi, Liverpool John Moores University, UK
Woontack Woo, GIST, Korea

Publicity Co-chairs

Fong-Lok Lee, CUHK, China
Jorge Posada, VICOMTech, Spain
Gloria Brown Simmons, University of California – Irvine, USA

Publication Chairs

Ming-Yong Pang, Nanjing Normal University, China

Financial Co-chairs

Huimin Shi, Nanjing Normal University, China
Mingmin Zhang, Zhejiang University, China

Organizing Co-chairs

Yi Li, Nanjing Normal University, China
Zhengxin Sun, Nanjing University, China
Jun Feng, Hohai University, China

General Secretary

Ruwei Yun, Nanjing Normal University, China

Executive Committee

Zhigeng Pan, Zhejiang University, China
Ruth Aylett, Heriot-Watt University, UK
Jim Chen, George Mason University, USA
Hyun-seng Yang, KAIST, Korea
Yi Li, Nanjing Normal University, China
Xiaopeng Zhang, CAS Institute of Automation, China
Adrian David Cheok, NUS, Singapore
Ryohei Nakatsu, Kwansei Gakuin University, Japan

Program Committee

Maha Abdallah, *France*
Sangchul Ahn, *Korea*
Isabel M. Alexandre, *Portugal*
Elisabeth Andre, *Germany*
Dominique Archambault, *France*
Ruth Aylett, *UK*
Leandro Balladares, *Mexico*
Christian Bauckhage, *Germany*
Rafael Bidarra, The *Netherlands*
Paul Brna, *UK*
Gloria Brown-Simmons, *USA*
Eliya Buyukkaya, *France*
Yiyu Cai, *Singapore*
Christophe Chaillou, *France*
Tak-Wai Chan, *Taiwan*, *China*
Maiga Chang, *Canada*
Yam San Chee, *Singapore*
Jim X. Chen, *USA*
Nuno Correia, *Portugal*
Jinshi Cui, *China*
Akshay Darbari, *India*
stephane Donikian, *France*
Stephen Downes, *Canada*
Abdennour El Rhalibi, *UK*
Guangzheng Fei, *China*
Paul Fergus, *UK*
Bernd Froehlich, *Germany*
Marco Furini, *Italy*
Lisa Gjedde, *Denmark*
Martin Goebel, *Germany*

Michael Haller, *Austria*
Wenfeng Hu, *China*
KinChuen Hui, *Hong Kong*, *China*
Wijnand IJsselsteijn, *The Netherlands*
Manjunath Iyer, *India*
Marc Jaeger, *France*
Posada Jorge, *Spain*
Bill Kapralos, *Canada*
Börje Karlsson, *Brazil*
Mike Katchabaw, *Canada*
Gerard Jounghyun Kim, *Korea*
Sehwan Kim, *USA*
Mona Laroussi, *France*
Fong-Lok Lee, *Hong Kong*, *China*
Sang-goog Lee, *Korea*
Jongweon Lee, *Korea*
Xin Li, *USA*
Craig Lindley, *Sweden*
Kwan-Liu Ma, *USA*
Bruce Maxim, *USA*
Wolfgang Mueller, *Germany*
Stéphane Natkin, *France*
Yoshihiro Okada, *Japan*
Chunhong Pan, *China*
Zhigeng Pan, *China*
Mingyong Pang, *China*
Jun Park, *Korea*
Daniël Pletinckx, *Belgium*
Edmond C. Prakash, *UK*
Marc Price, *UK*
Matthias Rauterberg, *The Netherlands*
Theresa-Marie Rhyne, *USA*
Albert Skip Rizzo, *USA*
Marco Roccetti, *Italy*
Cecilia Sik Lanyi, *Hungary*
Ulrike Spierling, *Germany*
Gerard Subsol, *France*
Mohd Shahrizal Sunar, *Malaysia*
Kim Hua Tan, *UK*
Ruck Thawonmas, *Japan*
Harold Thwaites, *Malaysia*
Jaap Van Den Herik, *The Netherlands*
Frederic Vexo, *Switzerland*
Yangsheng Wang, *China*
Charlie Wang, *Hong Kong*, *China*
Toyohide Watanabe, *Japan*
Kevin Kok-Wai Wong, *Australia*

Woontack Woo, *Korea*
Hongping Yan, *France*
Michael Young, *USA*
Ruwei Yun, *China*
Xiaopeng Zhang, *China*

Table of Contents

E-Learning Platforms and Tools

E-Learning System for Education

Application of E-Learning Systems

E-Learning Resource Management

Interaction in Game and Education

Integration of Game and Education

Game Design and Development

Virtual Characters, Animation and Navigation

Graphics Rendering and Digital Media

Geometric Modeling in Games and Virtual Reality

WRITE: Writing Revision Instrument for Teaching English

Jia-Jiunn Lo[1], Ying-Chieh Wang[1], and Shiou-Wen Yeh[2]

[1] Department of Information Management, Chung-Hua University Taiwan, Republic of China
{jlo,m09510033}@chu.edu.tw
[2] Department of Applied Linguistics and Language Studies, Chung-Yuan Christian University, Taiwan, Republic of China
shiouwen@cycu.edu.tw

Abstract. Corrective feedback and error correction are important tasks for ESL/EFL (English as a Second Language/English as a Foreign Language) writing instruction. Research findings showed that students' major difficulty in error correction lies in their failure to detect errors. Also, researchers proposed that error analysis can be reinvented in the form of computer-aided error analysis, a new type of computer corpus annotation. Annotations on digital documents can be easily shared among groups of people, making them valuable for a wide variety of tasks, including providing feedback. This study developed a web-based online corrective feedback and error analysis system called WRITE (Writing Revision Instrument for Teaching English). With this system, teachers can make error corrections on digitized documents, on the general web browser such as Microsoft Internet Explorer, with online annotations in the same way as the traditional paper-based correction approach. The WRITE system can feedback correct answers, teachers' comments, and the grammatical error type for each error to students. In addition, this system can provide users the annotation marks subject to different query conditions so that the problem of cognitive overload can be avoided. For error analysis purposes, this system can access the database and analyzes students' errors and displays the results as requested. Students use WRITE will be able to effectively identify more errors. Moreover, the ways that the corrective feedback delivered through the online annotation system can be used by students to develop his/her corrective strategies.

Keywords: online annotation, error correction, error feedback, error analysis, computer assisted language learning (CALL), writing instruction.

1 Introduction

Writing processes include tasks such as planning, transcribing, and revising (Ogata et al., 1999). Revision is often defined as the last stage in and the heart of the writing process. However, it is not an easy task. To most students, revision means correction (Lehr, 1995). As cited by Lehr, Adams (1991) proposed: *"Merely requiring students to revise or just spend more time revising will not necessarily produce improved writing."*

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 1–8, 2008.

Therefore it is critical for ESL/EFL (English as a Second Language/English as a Foreign Language) teachers and learners to receive a more constructive approach and a more interactive environment for corrective feedback and error correction.

Corrective feedback is a technique to help learners correct errors by providing them with some kind of prompting. As defined by Ellis (2007), corrective feedback takes the form of responses to text or utterances containing an error. The responses can consist of (1) an indication that an error has been committed, or (2) provision of the correct target language form, or (3) metalinguistic information about the error, or any combination of these. Corrective feedback is an area that bridges the concerns of teachers, researchers, and instructional designers. Although it is generally agreed that students expect teachers to correct written errors and teachers are willing to provide them, the immediate concern of many teachers "is not so much *to correct* or *not to correct*", but rather when and how to respond to what students write (Lee, 2003).

Much research has been conducted to search for effective writing feedback and correction methods. In responding to the limitations of paper-based error feedback and analysis, researchers have suggested a more constructivist approach to designing open-ended learning environments. Teachers should consider new and emerging technologies and the capabilities they add to approaches for teaching and supporting the distant learner (Ware & Warschauer, 2006). From the perspective of instructional design, traditional paper-based error feedback and analysis can be reinvented in the form of computer-aided error analysis, which is a potential type of computer corpus annotation.

Annotations are the notes a reader makes to himself/herself, such as students make when reading texts or researchers create when noting references they plan to search (Wolfe, 2002). Annotations are also a natural way to record comments and ideas in specific contexts within a document. Annotation systems can take advantage of networked technologies to allow communities of readers to comment on the same virtual copy of a text (Yeh et al., 2006). Compared to paper-based annotations shared merely through printed technology, online annotations provide readers with more opportunities for dialogue and learning through conversations (Wolfe, 2002). Annotations on digital documents are easily shared among groups of people, making them valuable for a wide variety of tasks, including providing feedback. As a language learning tool, online annotations for ESL/EFL writing seems to fit with the current trend of distance learning, cognitive conditions for instructed second language acquisition (Skehan, 1998). This study proposes that traditional paper-based corrective feedback and error correction method for EFL/ESL writing instruction can be reinvented in the form of computer-mediated corrective feedback and error correction using online annotation technology.

Many instructors have recently advocated the benefits annotations might have for developing language learners (Ogata et al., 1999). Practically, online annotations can be quite useful, in which students could share their annotations to discuss reactions to a text, or they could use annotations as a type of reading journal to share with the instructor. Basically, online annotations can provide a good way for writers to share knowledge and allow extended conversations to take place in the context of a common text. By facilitating easy movement between texts, annotation tools can emphasize the intertextual nature of reading. Tools for manipulating and rearranging annotations can scaffold different information strategies that help students learn to move from reading to writing. Also, as Bargeron, et al, (1999) claimed, annotations can provide "in

context" personal notes and can enable asynchronous collaboration among groups of users. With annotations, users are no longer limited to viewing content passively on the web, but are free to add and share commentary and links, thus transforming the web into an interactive medium.

However, in spite of the advantages mentioned above, the question of how annotations may help students' writing has not been sufficiently addressed (Wolfe, 2002). Studies investigating using annotation systems in ESL/EFL error feedback and analysis are especially needed. Based on the above discussion, this study develops an online annotation system, called WRITE (Writing Revision Instrument for Teaching English), which can provide annotation analysis and knowledge sharing, and can be applied to error correction, error feedback, and error analysis in English writing instruction.

2 The WRITE System

The WRITE system is based on the client/server architecture as illustrated in Fig. 1.

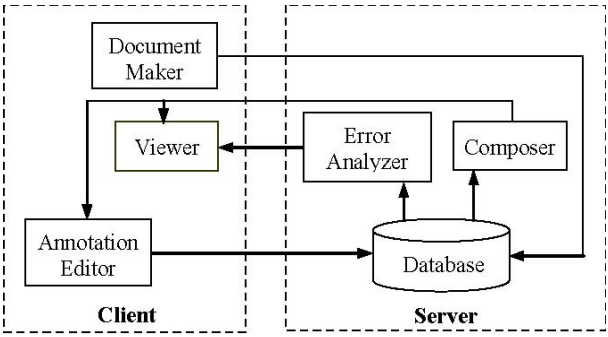


Fig. 1. The WRITE System Architecture

2.1 Document Maker

Document Maker is where students input their documents. As the document is edited, the system will convert it into the HTML format and save it in Document Database so that it can be displayed with general web page browsers for error correction marking by teachers.

2.2 Annotation Editor

Annotation Editor (Fig. 2) is where teachers input their correction markings of the document. It is implemented on the general web browser such as Microsoft Internet Explorer. In Annotation Editor, teachers can make correction marks and comments only, i.e., it is under "read-only" status in that the content of the original document cannot be changed. Such functionality of making correction marks under "read-only" status is quite important for students to be able to easily compare their original works and the corrective feedback.

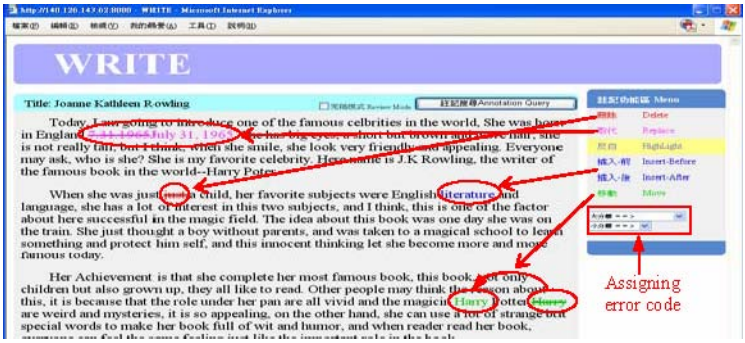


Fig. 2. Illustration of Annotation Editor (Annotation Mode)

To create a correction and comment, the teacher first highlights the text, named *annotation keywords*, to which he/she wants to annotate. Then the teacher assigns an error code by using two pull-down menus to indicate its major error category and error type. In WRITE, five major error categories are applied: (1) writing style, (2) composition structure, (3) sentences, (4) words and phrases, and (5) agreement, tense, and voice. Under each major error category, there are different numbers of error types (Yeh et al., 2006). After assigning the error type, the teacher clicks on one of the annotation tools to activate the corresponding function to place the error correction mark into the annotation keywords. The annotation tools include “Delete”, “Replace”, “HighLight”, “Insert-Before”, “Insert-After”, and “Move”. Then the WRITE system will use JavaScript to automatically insert the <SPAN> tag around the annotation keywords for showing the effects of annotation marks and store all related annotation information in Annotation Database. As the teacher moves the cursor over the annotation mark, related annotation information will be shown and the teacher can delete the correction mark by clicking the “Delete this annotation” button (Fig. 3).

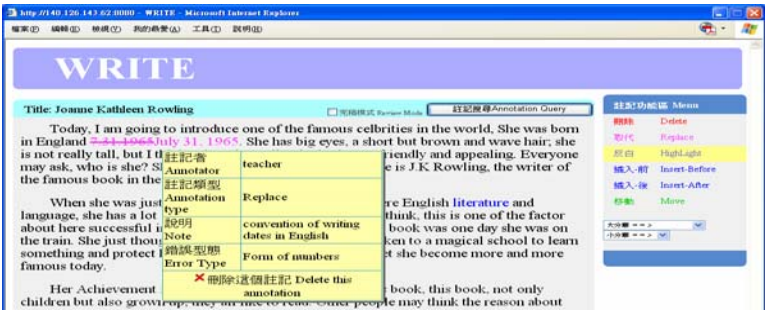


Fig. 3. Illustration of annotation information

One of the innovative functionalities of WRITE is the teacher can freely switch between the *annotation mode* (see Fig. 2) and the *review mode* (Fig. 4) to neatly review the “right” document after correction without showing the correction marks (*cf.* Fig. 2). In review mode, the annotation tools are hidden.

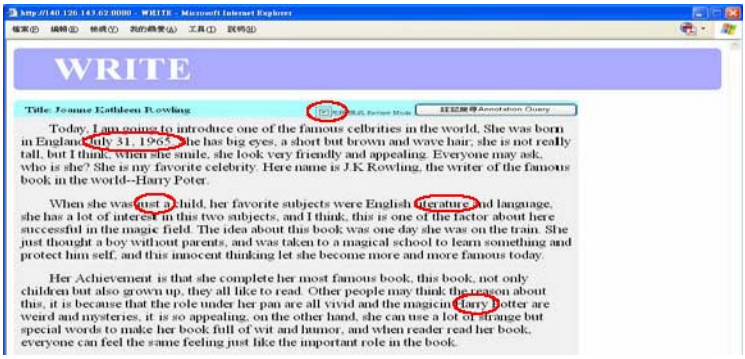


Fig. 4. Illustration of Annotation Editor (Review Mode)

Anchoring annotation positions is challenging in digital document annotation. The problem of "orphan annotations" is one of the major complains of annotation systems (Brush, 2002). In WRITE, since the <SPAN> tags of the annotation marks are inserted around the annotation keywords, the problem of anchoring annotation positions can be avoided even when the document is modified (Fig. 5).

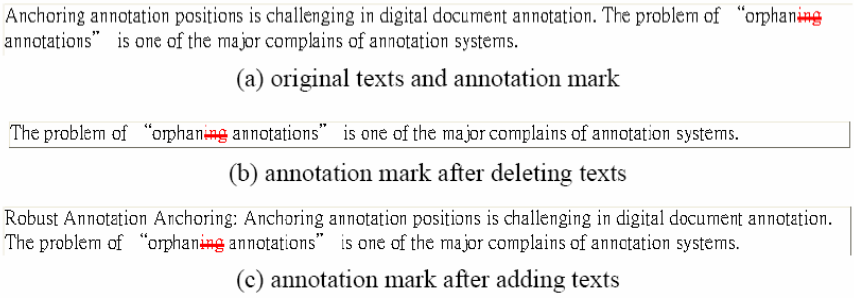


Fig. 5. Illustration of Robust Annotation Anchoring

2.3 Database

Two database modules are included in the system, Document Database and Annotation Database. Document Database stores the documents students written in HTML format through Document Maker. Annotation Database stores the related information of annotations, such as annotator, annotation type, error type, annotation identification code, annotation notes, etc. In WRITE, a unique annotation identification code is assigned to each annotation. It can make dynamic control to the annotation keywords by regarding each annotation as an object stored in the annotation database. Annotation Database offers the information for annotation query (manipulated by Composer) and error analysis (manipulated by Error Analyzer).

2.4 Composer

Since a document can be annotated with different annotation tools and error types, students might be confused due to too much information. Through Composer, with Annotation Database and the annotation identification code, the WRITE system can provide users the annotation marks subject to different query contitions so that the problem of cognitive overload can be avoided. In WRITE, a user can query the annotations based on the annotation type and the error type. In addition, the WRITE system can implement full-text search within "Annotation keyword". "Replaced words", and/or "Annotation notes" (Fig. 6).

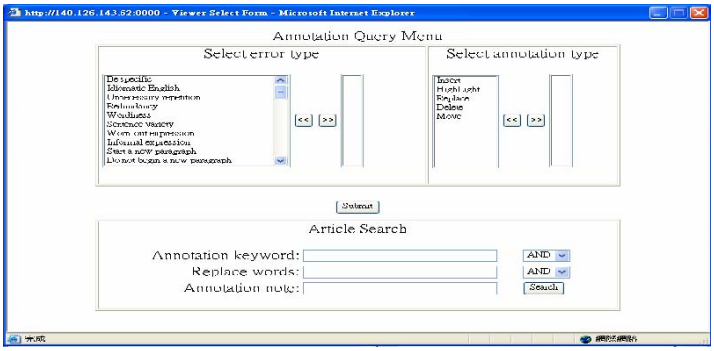


Fig. 6. Illustration of Annotation Query Menu

2.5 Error Analyzer

For error analysis purposes, Error Analyzer accesses the database and analyzes students' errors to display the statistical results of student error distributions in bar charts as requested by the teacher. Four error statistical analysis options are included:

Fig. 7. Illustration of Analysis Result of Single Document for an Individual Student (Analyzed Result Viewer)

single document for an individual student, all documents for an individual student, single document for a group of students, and all documents for a group of students (Fig. 7). Error analysis of single document and all documents for an individual student is helpful to realize the most severe barrier a student faces in writing a particular document and the overall most severe barrier a student faces in writing. On the other hand, error analysis of single document and all documents for a group of students is helpful to realize the errors most students have made in writing a particular document and the overall errors most student faces in writing.

2.6 Viewer

Two Viewers are included in the system, Document Viewer and Analyzed Result Viewer. Document Viewer is where students can view their documents after being corrected by the teacher. Like Annotation Editor, the student can freely switch between the *annotation mode* to view correction marks and the *review mode* to neatly review the "right" document without showing the correction marks. It is different from the Annotation Editor (see Figure 2) in that the annotation tools are hidden for both annotation mode and review mode. Through Document Viewer, students can know which parts of their documents are corrected and get detailed error feedback by moving the cursor over the annotation marks (see Fig. 3). Analyzed Result Viewer displays the four error analysis options analyzed by Error Analyzer (see Fig. 7).

3 Conclusions

Corrective feedback and error correction are important tasks for ESL/EFL writing instruction. EFL/ESL learners have great diversities in error correction and feedback strategies, and a more constructive approach and a more interactive environment for error feedback and error correction are needed. Research findings showed that students' major difficulty in error correction lies in their failure to detect errors. Also, researchers proposed that error analysis can be reinvented in the form of computer-aided error analysis, a new type of computer corpus annotation. Annotations on digital documents are easily shared among groups of people, making them valuable for a wide variety of tasks, including providing feedback. This study developed a web-based online corrective feedback and error analysis system called WRITE. With this system, teachers can make error corrections on digitized documents with online annotations in the same way as the traditional paper-based correction approach. The WRITE system can feedback correct answers, teachers' comments, and the grammatical error type for each error to students. In addition, this system can provide users the annotation marks subject to different query conditions so that the problem of cognitive overload can be avoided. For error analysis purposes, this system can access the database and analyzes students' errors and displays the results as requested. Four error analysis options are included: single document for an individual student, all documents for an individual student, single document for a group of students, and all documents for a group of students. Students use WRITE will be able to effectively identify more errors; moreover, the ways that the corrective feedback delivered through the online annotation system can be used by the student writer to develop his/her

corrective strategies. However, future research is needed to confirm this hypothesis. Future research should also investigate the long-term effects of online annotations on student writing development.

Acknowledgements

This is part of a larger study that has been generously supported by the National Science Council of Taiwan, R.O.C. (NSC 96-2411-H-033-006-MY3).

References

1. Bargeron, D., Gupta, A., Grudin, J., Sanocki, E.: Annotations for Streaming Video on the Web: System Design and Usage Studies. WWW 8, Toronto, Canada, 61–75 (1999)
2. Brush, A.: Annotating Digital Documents for Asynchronous Collaboration. Technical report (2002)
3. Ellis, R.: Corrective Feedback in Theory, Research and Practice. In: The 5th International Conference on ELT in China & the 1st Congress of Chinese Applied Linguistics, May 17-20, Beijing Foreign Language Studies University, Beijing, China (2007), Retrieved October 23, 2007 from the World Wide Web: `http://www.celea.org.cn/2007/edefault.asp`
4. Lee, I.: L2 writing teachers' perspectives, practices and problems regarding error feedback. Assessing Writing 8, 216–237 (2003)
5. Lehr, F.: Revision in the Writing Process. Accessed online (2007/11/6): `http://www.readingrockets.org/article/270`
6. Ogata, H., Feng, C., Hada, Y., Yano, Y.: Computer Supported Proofreading Exercise in a Networked Writing Classroom. In: ICCE 1999 (1999)
7. Skehan, P.: A Cognitive Approach to Language Learning. Oxford University Press, Oxford (1998)
8. Ware, P.D., Warschauer, M.: Electronic feedback and second language writing. In: Hyland, K., Hyland, F. (eds.) Feedback in Second Language Writing: Contexts and Issues, pp. 118–122. Cambridge University Press, New York (2006)
9. Wolfe, J.: Annotation technologies: A software and research review. Computers and Composition 19, 471–497 (2002)
10. Yeh, S.-W., Lo, J.-J., Huang, J.-J.: The Development of an Online Annotation System for EFL Writing with Error Feedback and Error Analysis. In: ED-MEDIA 2006, Orlando, Florida, USA, pp. 2480–2485 (2006)

u-Teacher: Ubiquitous Learning Approach

Zacarías F. Fernando[1], Cuapa C. Rosalba[2], Lozano T. Francisco[3], Vazquez F. Andres[4], and Zacarías F. Dionicio[5]

Benemérita Universidad Autónoma de Puebla,
[1,3,4]Computer Science and [2]TCU-111 and [5]Mathematics
14 Sur y Av. San Claudio, Puebla, Pue.
72000 México
[1]fzflores@yahoo.com.mx, [2]rcuapa_canto@yahoo.com
[3]pakokonka@hotmail.com, [4]andrexsol@gmail.com,
[5]jzacarias@fismat1.fcfm.buap.mx

Abstract. The Learning of the century XXI demand ubiquitous characteristics that allow to learner not only to have information available in any place, in any time and in any way. But rather, to have the information in right time, in the right place and in right way [3]. An ubiquitous learning environment should allow the learner to have these three characteristics at your disposal in all moment. Considering that a lot of people have access to cell phone technology and it would be fantastic to use this technology as a learning tool. We have developed an ubiquitous learning system (on cellular phones) that also it combines web technologies and cell phone. Our system shows that this novel form of learning is completely accepted and it increases the learning level significantly.

Keywords: Learning, Ubiquitous, Learner, PDA, Cellular Phone.

1 Introduction

Mobile devices are part of our everyday environment and consequently part of our educational landscape [9]. The current mobile trends in education have demonstrated that learning no longer needs to be classroom. Current trends suggest that the following three areas are likely to lead the mobile movement: m-learning, e-learning and u-learning. There are estimated to be 2.5 billion mobile phones in the world today. This means that this is more than four times the number of personal computers (PCs), and todays most sophisticated phones have the processing power of a mid-1990s PC. Even, in a special way, many educators are already using iPod in their curricula with great results. They are integrating audio and video content including speeches, interviews, artwork, music, and photos to bring lessons to life. Many current developments, just as ours, incorporate multimedia applications. We allow to educators and students can create their own content. Therefore, its a great way for educators to create, organize, and distribute content. In the late 1980's, a researcher at Xerox PARC named Mark Weiser [8], coined the term "Ubiquitous Computing". It refers to

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 9–20, 2008.

the process of seamlessly integrating computers into the physical world. Ubiquitous computing includes computer technology found in microprocessors, mobile phones, digital cameras and other devices. All of which add new and exciting dimensions to learning.

The main characteristics of ubiquitous learning are shown as follows [1,2]:

- Permanency: Learners can never lose their work unless it is purposefully deleted. In addition, all the learning processes are recorded continuously in everyday.
- Accessibility: Learners have access to their documents, data, or videos from anywhere. That information is provided based on their requests. Therefore, the learning involved is self-directed.
- Immediacy: Wherever learners are, they can get any information immediately. Therefore learners can solve problems quickly. Otherwise, the learner may record the questions and look for the answer later.
- Interactivity: Learners can interact with experts, teachers, or peers in the form of synchronies or asynchronous communication. Hence, the experts are more reachable and the knowledge is more available.
- Situating of instructional activities: The learning could be embedded in our daily life. The problems encountered as well as the knowledge required are all presented in the nature and authentic forms. It helps learners notice the features of problem situations that make particular actions relevant.
- Adaptability: Learners can get the right information at the right place with the right way.

Moreover, ubiquitous learning can be Computer Supported Collaborative Learning (CSCL) environments that focus on the socio-cognitive process of social knowledge building and sharing. Therefore, in this paper we propose as objective a system of mobile learning, exploding for this the cellular phone technology, just as it is described in this document.

Our paper is structured as follows: In section 2 we describe the used development platform and general architecture of u-Teacher. Next, in section 3 we present the ubiquitous learning environment and the way in that learners interact in this. Section 4 contains the tasks designed for u-Teacher. Also, in section 5 we show the actions required to increase learning. The section 6 discuss the obtained results. Finally, the conclusions are drawn in section 7.

2 General Architecture of u-Teacher

In this section we present our general framework where our system was implemented, i.e., infrastructure development called NetBeans Mobility. Likewise, we present the general architecture of our system.

2.1 The NetBeans Mobility

In the development of our application we have used The NetBeans Mobility Packs as support for the two base configurations of the Java ME platform, CLDC and

CDC. The Connected, Limited Device Configuration (CLDC) is for small wireless devices with intermittent network connections, like mobile phones, and personal digital assistants (PDAs). The Mobile Information Device Profile (MIDP), which is based on CLDC, was the first finished profile and thus the first finished Java ME application environment. MIDP-compliant devices are widely available worldwide.

On the other hand, The Connected Device Configuration (CDC) is for larger devices (in terms of memory and processing power) with robust network connections, such as set-top boxes, Internet appliances, and embedded servers. The NetBeans IDE provides a wizard that enables you to quickly create a MIDP project. When creating the project, you can choose to develop your application in the Visual Mobile Designer (VMD) or in the Source Code Editor. The visual mobile designer's use allows you the ability to graphically plan out the flow of the application and design the screens the application will use. The designer automatically creates the code for the application.

2.2 Architecture of u-Teacher

The general architecture of u-Teacher (Fig. 1) consists basically of the following components:

- The applications server
- The Web Server
- Cell phone to attend multimedia messages
- Service of communications for mobile devices
- Firewall to reject intruders
- Intelligent agent to evaluate all test that learners send

First, the [illegible] contains our u-Teacher and allows that learner can download this system to his mobile device (allowing this way the u-Teacher multiplicity). Likewise it allows to maintain the application to service of learners through the web.

Second, the [illegible] is the responsible for the operation of u-Teacher. This provides two functions: first of all, to validate uploads and downloads from the site. Secondly, to maintain a permanent coordination with the cell phone.

Third, our system has a [illegible] to attend all requests via [illegible]. For instance, when learner has download a test to its mobile and it solves it off-line, then he can send their answers through mobile via a message.

Fourth, [illegible], Which has a firewall to prevent the entry of intruders in our system. Also, it allows the interaction among learners in their practices of their activities in the learning of the English language.

Fifth, [illegible], it is maintained by the department of communications of BUAP. However, it is vital for the correct operation of our application. Also, it is important to note that communications in mobile devices and security in these is provided by the company that provides the service.

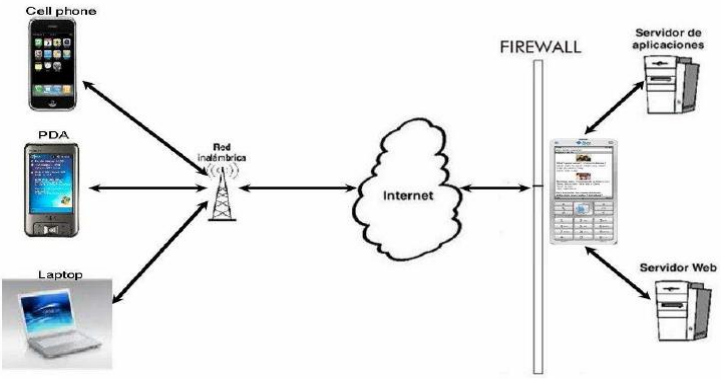


Fig. 1. General Architecture for u-Teacher

Finally, our application has an [illegible] whose tasks are: to evaluate the answers of the test either received through the web or, through messages via cell phone. As second task this the supervision of the correct update of new lessons, exams, homeworks, etc. The integration of all these features make our methodology doesnt just work, it works well! This ensures the efficiency of our system in the ubiquitous learning, as shown in the rest of paper.

3 The Ubiquitous Learning Environment

To design a learning environment ubiquitous is necessary to allow that each student interacts with many embedded devices. Just as Jones and Jo mentions it [5], this relationship is common in the evolving ubiquitous computing era. In the ubiquitous classroom, students move around ubiquitous space (u-space) and interact with the various devices.

In the Autonomous University of Puebla we are in a transformation process in the teaching [7]. One line of them corresponds exactly to the learning based on new technologies. Here, each student will carry a wireless device PDA or mobile phone. u-Teacher will be available in all moment and in anyone of the wireless technologies (PDA, cell phone or Laptop). Besides, u-Teacher allows learners to interact in two ways: off-line and on-line. On-line through the Web and, Off-line through cell phone or PDA without any kind of connection to the internet. When learner works on-line makes it through Internet. If on the other hand, makes off-line, communication through cell phone is via multimedia messaging. Also, learner can download lessons, homeworks and exams to your mobile cell phone to be able to study off-line in anywhere and anytime. This Characteristic has allowed our students to interact more with each other and their teachers. So, our teachers come to school every day with the aim of not only instructing their students in their academic subjects, but also to develop in them an enthusiasm for learning throughout their lives and the new technologies as part of their daily lives. With this new proposal we have achieved that our learners see in

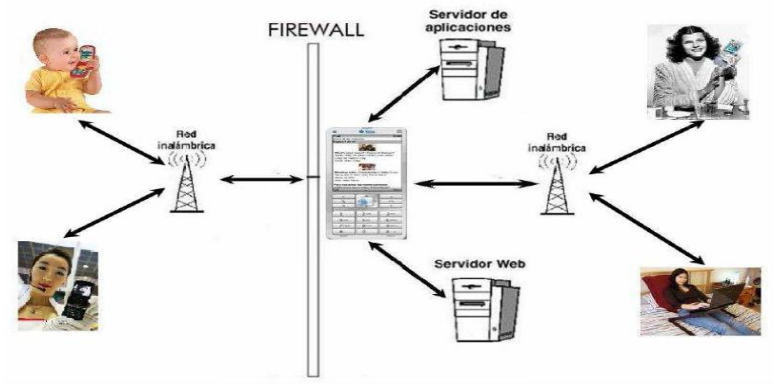


Fig. 2. Interaction of Learners into u-space

the learning to an ally. They do this by making learning an exciting experience, one that focuses on seeking fundamental knowledge rather than aspiring to the highest score on an exam.

Learning theories are important in the design of educational technology because it helps to create a relationship among the information, the learner, and the environment [4]. For this reason, as we can observe in Fig. 2 we have included pedagogical information which is based on constructivist theory, allowing students to create knowledge from what they see, hear, read and perceive. You can see it every day in every classroom and home. Even a casual observer will recognize that our learners are fully engaged in learning. This is because our system is based on one of the technologies most accepted by the whole world. They schedule their time, arrange their days, and complete their assigned and selected works. Through these accomplishments, they develop a high level of self-discipline, responsibility and maturity. Furthermore, the ubiquitous learning environment is a situation or setting of pervasive (or omnipresent) education (or learning). Education is happening all around the student but the student may not even be conscious of the learning process. The use of wireless and mobile technology makes them easily accessible and contributes to educational functionality. The wireless and mobile devices include mobile phones, PDAs and Laptops. Our proposal contains all the new technologies available, and the education should be offered in anytime, anywhere and anyway.

4 u-Teacher Model

Our proposal is the combination of mobile technologies and web technologies. In Fig. 3 we can appreciate the main interface our application and it is important to point out that all interfaces developed for the web are also available for the cell phone. Even, we can assert that this is the feature that has motivated and increased the level of learning in our students.

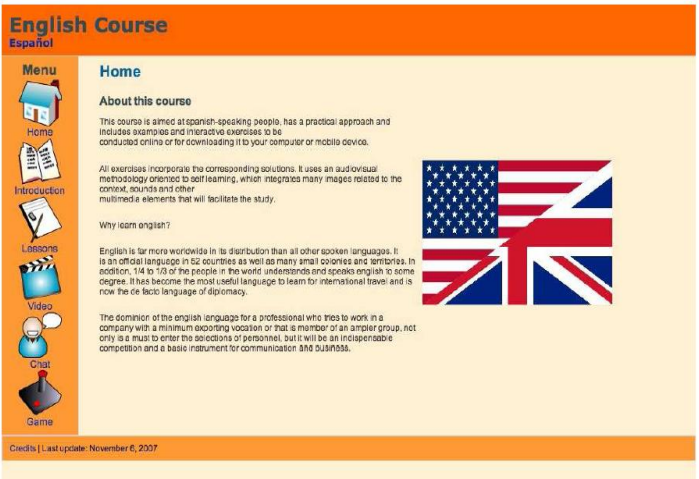


Fig. 3. Main interface

In this system we have combined different characteristics that make an avant-garde proposal in the development of novel learning technologies. We have developed a system based on ubiquitous learning. This proposal adds the component called "ubiquitous space" (u-space). This concep allows to learners to interact with ubiquitous objects/devices. Each student is part of the many to one relationship within this u-space. It is immaterial which particular device the student is currently interacting with, as all devices are networked and communicating within the Ubiquitous Space.

The tasks designed for this system are:

- Download a set of lessons to help students learn the basics of English language online or in their mobile devices (offline).
- Basic sentences syntax (text and voice).

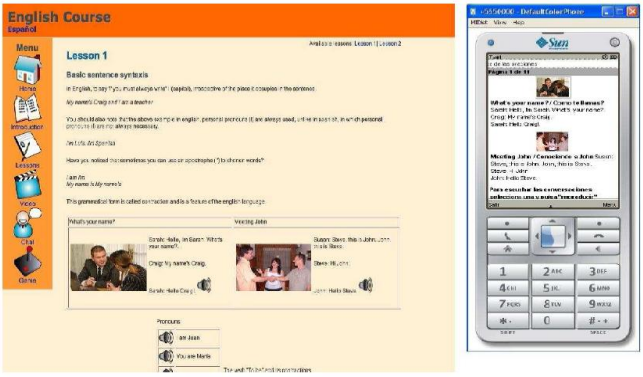


Fig. 4. Lessons that include video, audio and test

- Exercises of pronunciation students can listen to the correct pronunciation.
- Pronunciation exercises where students can record their pronunciation, and if possible, compare it with the correct one.
- Videos where the user watches and listens to brief dialogues.
- Formats that validate the understanding of reviewed videos.
- Collaboration practices where at the most 4 students can interact with questions and answers.
- Establishment of dialogues and corresponding activities for them.
- Games that help to the learning.
- Chat (text and voice) via Bluetooth between devices that allow the learning of this languages writing.

5 Learning Requires Action

Learning English requires action. In Fig. 4 and Fig. 5 we can observe that the actions to realize are very important to learn English. Furthermore, the fact that the actions or tasks can carry out them in both portal or in cell phone it is an added value. You may know all the learning tips, but if you don't start doing things, you will achieve nothing. The fact is if you want to learn to speak English well, you must change your life. Some examples of things you will have to do:

- Read a book in English for an hour every day, analyzing the grammar in sentences and looking up words in an English dictionary.
- Listen to an audio-book or other recording in English, stopping it frequently, trying to understand what is being said, and trying to imitate the speaker's pronunciation.
- Spend your afternoon practicing the pronunciation of the English "r" sound.
- Carefully write an e-mail message in English, using a dictionary or a Web search every 20 seconds to make sure every word is correct, and taking 5 minutes to write one sentence.
- Think about an English sentence you've read, wondering if it could say "a" instead of "the" in the sentence, and trying to find similar sentences on the Web to find out the answer.
- Walk down the street and build simple English sentences in your head (talking to yourself in English about the things you see around you).

By virtue of these actions we have implemented lessons that cover these aspects (Fig. 4 and Fig. 5). These characteristics have proven to give excellent results in the use of our tool. these results obey to that learners could make use of his course of English in anywhere and anytime. Our tool has been considered as part of a university initiative to encourage creative uses of technology in education and campus life. The problem with learning and teaching English as a foreign language is that all English learners want to speak English well; however, most learners don't want to spend time on learning English on their own. (Why they sign up for English classes and hope their teacher will force knowledge into their heads?.) This lack of motivation means that learners basically don't spend

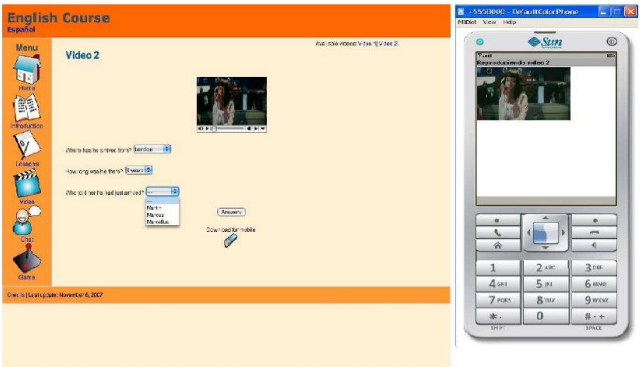


Fig. 5. Interactive videos for listening

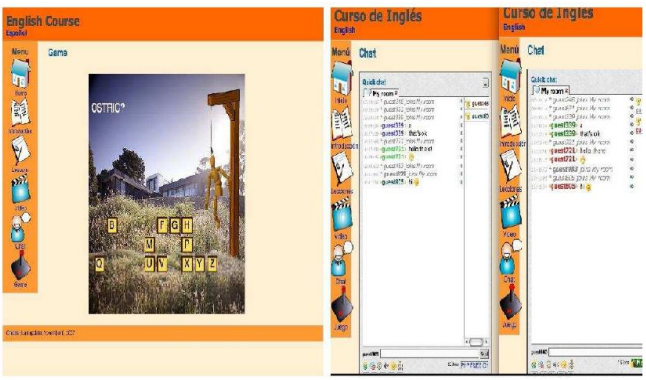


Fig. 6. Interactive games for learning

their own time on learning English, and if they do, they don't do it regularly. For example, a typical learner might study English phrasal verbs for 12 hours before an English exam. However, he will not read a book in English for 30 minutes every day. He just doesn't feel that learning English is pleasant enough, so he will only do it if he has to. The problem is that a huge one-time effort gives you nothing, while small, everyday activities will give you a lot. If you are one of those learners and don't feel like practicing the pronunciation of the "r" sound or thinking about English sentences every day, we have news for you: You're going to have to make yourself want to do these things. In other words, you'll have to work on your motivation. Fortunately, there are proven techniques to help you with that.

Some other basic things we must remember to take into account when learning another language are:

- Motivation: Become a person who likes to learn another language.
- Dictionary: Get a good dictionary.

- No mistakes: Avoid mistakes. Try to use the correct form of the language from the beginning.
- Pronunciation: Learn to pronounce that language sounds. Learn to understand phonetic transcription and the phonetic alphabet.
- Input: Get the language into your head by reading and listening to lots of sentences in that language, you could read or watch movies.

Language is primarily a spoken form of communication. We learn our native language as children by hearing the spoken language and then imitating it. This is something often overlooked. I believe that the most successful language learning methods are more audio-based than otherwise. In fact, I learned English that way. We can't ignore the importance of reading, but clearly the most fundamental aspect of communicating in a language is speaking and listening.

Videos are another important part that supplements the actions that should be carried out (see Fig. 5). While more we listen conversations and sentences in that language more quick will advance in our learning. Videos, movies and music stimulate the auditory sense (such as in Fig. 4 and Fig. 5); this allows making an imitation of the pronunciation more clear and correct. This portal tends to improve the users knowledge of the English language throughout practicing with exercises that will be asked to complete. If the user has little or null knowledge of this language, he shall obtain the bases to be able to begin to speak it and to understand it. In particular, we have incorporated the reproduction of videos to our mobile devices with the objective of offering to learner the opportunity to practice in anytime and anywhere taking advantage of the mobile technology.

With respect to the test, we have implemented an agent whose task is the one of validating that the answers are correct. Here, the learner answers the test offline (anytime or anywhere) and then he chooses the upload option. The answers are sent to the portal through a file (text message) so that they are evaluated by our agent.

The result of the test evaluation is sent to the apprentice through a text message. With respect to Chat, this it is available through the interactive portal only. With this characteristic we provide to learner with a mobile tool that offers bigger educational opportunities. u-Teacher contains an amusing game that absorbs the learner in the subconscious learning of such an extensive vocabulary as it is desired. It is important to point out that in this first version we have added only the hangman game. However, due to the high index of interest shown by u-Teacher's users regarding this game, we are developing new games that the same as it happened with this they capture the attention of learners.

In [6] the authors describe five properties of handheld computers that produce unique educational affordances:

- Portability - can take the computer to different sites and move around within a location
- Social Interactivity - can exchange data and collaborate with other people face to face
- Context Sensitivity - can gather data unique to the current location, environment, and time, including both real and simulated data

- Connectivity - can connect handhelds to data collection devices, other handhelds, and to a common network that creates a true shared environment
- Individuality - can provide unique scaffolding that is customized to the individual's path of investigation.

6 Discussion on the Results

As expected, foreing language and video courses integrated the device, but its use also extended to other social science and humanities courses. In addition, all first-year engineering students used the cell phone in their foreing language course. Audio-intensive courses reported that the cell phone increased the frequency and depth of student interaction with audio course content through portable and flexible access offered by the cell phone, PDA's, and laptops. Initial planning for academic cell phone use focused on audio playback; however, digital recording capabilities ultimately generated the highest level of student and faculty interest. Both recording and listening were the most widely used feature for academic purposes, with 50% of first-year students reporting using the cell phones recording ability for academic purposes. This high level of interest in digital recording and listening were also reflected in the proposals received and supported. Another characteristic that has been sued is to incorporate a Karaoke for cell phone. This is because young people show a high interest in music in English.

Benefits of academic u-Teacher use

- Convenience for both faculty and students of portable digital course content, and reduced dependence on physical materials.
- Flexible location-independent access to digital multimedia course materials, including reduced dependence on lab or library locations and hours.
- Effective and easy-to-use tool for digital recording of interviews, field notes, small group discussions, and self-recording of oral assignments.
- Greater student engagement and interest in class discussions, labs, practices outside of class, and bigger concentration.
- Enhanced support for individual learning preferences and needs.

7 Conclusions

In this paper we have presented the impact that our called tool u-Teacher has had in the teaching of the English language. u-Teacher uses ubiquitous technology and the concept of ubiquitous learning. This is part of ongoing research and development being undertaken at Autonomous University of Puebla. The research team is extensively involved in ubiquitous technology and communications. Innovation and experimentation for academic u-Teacher use was widely reported as well, with 75% of firt-year students reporting having used at least one u-Teacher feutured in a class or for independent support of their studies. In

addition to the findings outlined above regarding academic u-Teacher use, the evaluation also identified some significant institutional impacts of the project:

- Increased collaboration and communication among campus technology support groups highlighted strengths and gaps in the existing technology environment and was an impetus for broader planning and improvement of infrastructure and services.
- To increase the development of this type of tools because these have been well accepted by the students.
- The new technologies evolve and more pervasive forms of technology emerge, computers will become "invisible" and will be embedded in all aspects of our life.
- Finally, u-Teacher has allowed us to increase the learning until in 80%. This has motivated the development from other similar applications to the one presented in this paper.

Acknowledgments

Thank you very much to the Autonomous University of Puebla for their financial support. This work was supported under project VIEP register number 15968. Also, we thank the support of the academic body: Sistemas de Informacin.

References

1. Chen, Y.S., Kao, T.C., Sheu, J.P., Chiang, C.Y.: A Mobile Scaffolding-Aid-Based Bird -Watching Learning System. In: Proceedings of IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE 2002), pp. 15–22. IEEE Computer Society Press, Los Alamitos (2002)
2. Curtis, M., Luchini, K., Bobrowsky, W., Quintana, C., Soloway, E.: Handheld Use in K-12: A Descriptive Account. In: Proceedings of IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE 2002), pp. 23–30. IEEE Computer Society Press, Los Alamitos (2002)
3. Fischer, G.: User Modeling in Human-Computer Interaction. Journal of User Modeling and User-Adapted Interaction (UMUAI) 11(1/2), 65–86 (2001)
4. Jacobs, M.: Situated Cognition: Learning and Knowledge Relates to Situated Cognition (1999) [verified 31 Oct 2004] http://www.gsu.edu/~mstswh/courses/it7000/papers/situated.htm
5. Jones, V., Jo, J.H.: Ubiquitous learning environment: An adaptive teaching system using ubiquitous technology. In: Atkinson, R., McBeath, C., Jonas-Dwyer, D., Phillips, R. (eds.) Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference, Perth, December 5-8, pp. 468–474 (2004), http://www.ascilite.org.au/conferences/perth04/procs/jones.html
6. Klopfer, Squire, Holland, Jenkins: Environmental Detectives-the development of an augmented reality platform for environmental simulations, Educational Technology Research and Development. Springer, Heidelberg (2002)

7. Zacarias, F.F., Lozano, T.F., Cuapa, C.R., Vazquez, F.A.: English's Teaching Based On New Technologies. The International Journal of Technology, Knowledge & Society, Northeastern University in Boston, Massachussetts, USA, Common Ground Publishing, USA (2008)
8. Weiser, M.: The computer for the twenty-first century, pp. 94–104. Scientific American (September 1991)
9. Zacarías, F., Sánchez, A., Zacarías, D., Méndez, A., Cuapa, R.: Financial Mobile System Based On Intelligent Agents in the Austrian Computer Society book series, Austria (2006)

A Model for Knowledge Innovation in Online Learning Community

Qinglong Zhan

Department of Computer Science, Tianjin University of Technology and Education, Tianjin 300222, China
qlzhan@126.com

Abstract. This paper describes how to introduce the mechanism of knowledge innovation into online learning community (OLC) and constructs a model which can facilitate innovative knowledge and develop online learner's ability of knowledge innovation. The model is represented by three levels, namely: individual, collaborative and intermediary level, which is based on theories of knowledge creation and management, cognitive and social constructivism. Individual knowledge innovation begins with internalization, via combination, externalization and socialization, which is different from Nonaka's SECI. In collaborative knowledge innovation, learners in OLC share, compare, negotiate, create and integrate knowledge together. Individual knowledge innovation and collaborative knowledge innovation need certain intermediary. In doing so, individual knowledge applied to OLC situation can produce, promote and create new knowledge of OLC.

Keywords: online learning community, knowledge innovation, SECI.

1 Introduction

With the development of e-learning, online learning community (OLC) plays a more important role in e-learning. OLC is not only the inevitable result of development of e-learning, but a basic component of it as well. Through utilizing modern information technology such as Internet that offers learning environments of whole new communication mechanisms and abundant resources, OLC is an important environment of knowledge innovation which can realize production, sharing, application and innovation of knowledge. Since 1990s, for the research on OLC, many researchers centralized on concept, mechanism of formation and growth, knowledge construction/building etc., but the research on knowledge innovation is an omitted paradigm.Therefore, this paper focuses on introducing mechanism of knowledge innovation into OLC and constructing a model which can facilitate innovative knowledge and develop online learner's ability of knowledge innovation.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 21–31, 2008.

2 Conceptual Background

2.1 OLC

OLC based on technologies of network and communication is an interactive autonomy cyberspace, in which several learners with common interest break through the space-time scope by web communication tools and study some course or theme for the goal of realizing knowledge sharing and knowledge innovation.OLC as a metaphor of learner group in traditional school, generated and formed gradually with the development of Internet, has broken through the boundary to limit on the basis of inheriting learner group's characteristics in traditional school, and possesses certain opening and freedom. So long as a learner who is interested in some course/theme joins in, through the exchange in certain time, forms a relatively stable online learner's group gradually, and becomes an OLC. OLC has some common interest and behavior codes, where each learner has the power and responsibility for participating in establishing and maintaining it. Learners facilitate their own learning and innovation of knowledge in OLC by sharing information, resources, mutual thoughts, views, artifacts and experiences.

2.2 Knowledge Innovation

According to scientific research perspective, knowledge innovation is a process of creating new knowledge in fundamental and technological sciences through scientific research, whose purposes are to pursue original new discoveries, to explore new laws, to create and provide original new theories and methods. Based upon daily perspective, the knowledge innovation need not create absolutely brand-new knowledge but renew and change knowledge, merge new elements or add new forms into existed knowledge, improve understandings and get new development on something existed, put forward new problem-solving, and so on. In knowledge management, knowledge innovation is the whole process of production, creation and application of knowledge.

2.3 Defining Knowledge Innovation in OLC

Combining the understanding of knowledge innovation with characteristics of OLC, I think knowledge innovation in OLC is the concept of daily perspective and a simple innovative activity, so there is a great difference from knowledge innovation activity in scientific research. Therefore, knowledge innovation in OLC is a process that individual online learner and community produce new viewpoints, thoughts and new problem solutions, change range and level of the existing knowledge structure, and eventually make new meaning through interaction and collaboration. Its key element is that new understanding emerges on the basis of original cognition. According to Bloom's classification of learning objectives, the knowledge which online learners produced through analyzing, synthesizing and appraising, belongs to knowledge of innovation.

3 OLC Is Regarded as the Environment of Knowledge Innovation

At present,both distance and classroom educational environments fetter the process of interactive knowledge innovation, but OLC as an innovative engine and "Ba" of Nonaka's term is one of the most effective environments in knowledge innovation and conversion. OLC as the sharing space where knowledge is created and innovated, can realize integration of virtuality (E-mail, online meeting) and intelligence (sharable experiences, thoughts and ideas).

3.1 OLC Forms the Foundation of Knowledge Innovation

OLC has common theme, goal, understanding, trust and open culture, which make up of the foundation of knowledge innovation. First of all, OLC learners always centre on the field of certain theme, and participate in the OLC because their learning contents and learning interests are closely related with the learning theme field. Therefore,while encountering problems about this theme, learners spontaneously get together to produce methods of problem-solving and form common innovative objectives. Secondly, OLC learners have questions of common concern, similar background and knowledge field, so they are apt to communicate between them and build the common understanding to the particular theme field. Thirdly, informality and opening in OLC which offers opportunities to freedom on informal dialogues and thought expression, can facilitate exchange and collaboration in an open atmosphere, "to achieve a deeper understanding of learning content and knowledge themes, to work together to solve problems, to exchange experience and develop new knowledge"(Seufert,2002).

3.2 The Mechanism in OLC Is Favorable to Knowledge Innovation

The mechanism of favorable innovation of knowledge in OLC is shown in the following several aspects.First, ambiguities of OLC boundary make information transmission and knowledge sharing more convenient; Second, OLC strengthens means and scope of communication when compared with traditional community; Third, being no formal institutional structure in OLC, learners can be freely carrying on exchange and collaboration on questions of common concern equally, make learners escape from institutional construction of traditional command and control, and turn to knowledge-intensive community which is more favorable to knowledge sharing. Hence, information and knowledge can be transmitted directly from a learner to any other learner in OLC without adopting the traditional exchange way, thus forming a knowledge community based on CSCL technologies.

3.3 Knowledge Sharing and Innovation Are Two Key Activities in OLC

Innovations arise at the intersection between flows of people and flows of knowledge (Starbuck, 1992). OLC, which emphasizes knowledge sharing through

network, facilitates the learner acquiring and sharing knowledge from other experienced learners by informal learning. It is an ideal environment of innovative process which is networking, interacting and knowledge-driving. So, knowledge sharing and innovation are two key activities in OLC. Knowledge innovation is promoted by knowledge sharing in which learners can obtain new elicitation, thinking or inspiration that provides the possibility of creating new knowledge.

3.4 OLC Strengthens the Process of Knowledge Integration

OLC can break through physical barriers and offer learners chances to share experiences and context knowledge, obtain enlightening knowledge from outside, make it possible to promote new understanding and explanation of knowledge, strengthen and absorb diversified new knowledge, and improve integration of new knowledge and existing knowledge.

3.5 OLC Influences the Process of Knowledge Utilization

Any kind of knowledge shared in OLC should be codified in someway in order to be digitized (Afuah, 2003), which facilitates knowledge memorization, retrieval and recombination (Fahey and Prusak, 1998). Digitization of knowledge can increase knowledge available in conversion, such as electronic files and knowledgebase-searching make it easy to discover, reorganize, externalize and internalize knowledge. Through allowing to obtain various knowledge in real time, OLC promotes learners to combine several conflictive-like knowledge into a new schema which will strengthen action and innovation. In OLC, learners not only can discern and contact more learners with different knowledge, but also allow their spontaneous connection to collaborate in developing concrete application of certain knowledge directly, create public knowledgebase and promote them to find the best application of their thoughts.

4 Construction of OLC Knowledge Innovation Model

Knowledge innovation in OLC is a systematized process, which includes two sub-processes of individual knowledge innovation and collaborative knowledge innovation. Individual knowledge innovation is the foundation of knowledge innovation of collaboration. In collaborative knowledge innovation, sharing, comparing, negotiating, creating and integrating knowledge together among the learners of OLC, produce new facts, understandings, concepts, viewpoints and theories. Individual knowledge innovation and collaborative knowledge innovation need certain intermediary, in doing so, individual knowledge applied to OLC situation can produce new knowledge and create and promote the knowledge of OLC. Therefore, OLC knowledge innovation model(Fig. 1)includes individual knowledge innovation layer, intermediary layer and collaborative knowledge innovation layer.

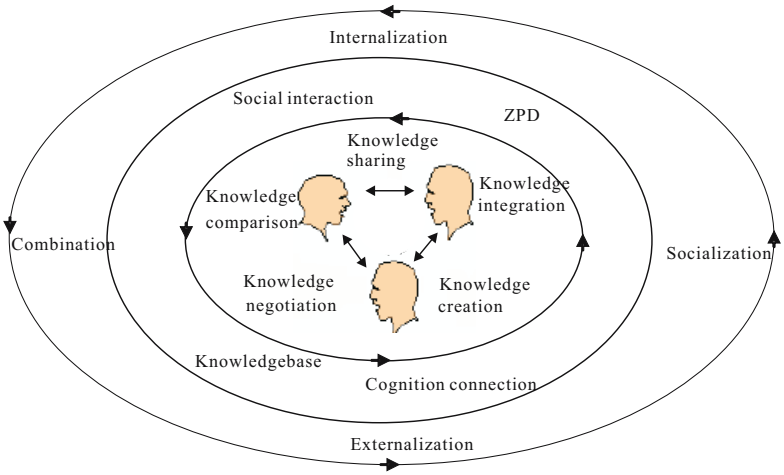


Fig. 1. A model of knowledge innovation in OLC

4.1 Individual Knowledge Innovation Layer in OLC

Any knowledge innovation begins with the individual. According to Nonaka, Takeuchi and Konno (2000), new knowledge is produced through continuous conversion between tacit knowledge and explicit knowledge. There are four modes of knowledge conversion. They are: (1) socialization, from tacit knowledge to tacit knowledge; (2) externalization, from tacit knowledge to explicit knowledge; (3) combination, from explicit knowledge to explicit knowledge; and (4) internalization, from explicit knowledge to tacit knowledge.

Internalization of individual knowledge. Internalization of individual knowledge is a process that a learner converts explicit knowledge in online course content and knowledgebase into tacit knowledge. By reading course materials, watching video, operating simulation, practising and testing online, searching knowledgebase, the learner forms individual internalized knowledge and experiences and builds personal knowledge foundation.

Combination of individual knowledge. That individual creatively uses technological tool, online course and knowledgebase OLC offered, can promote combination and conversion between explicit knowledge, produce new viewpoints and improve information processing. "An online program offers a structural configuration that meets the purpose of the course and the learners' needs"(Kearsley and Lynch, 1996).Knowledgebase helps individuals to reconfigure and create new knowledge of existing knowledge. A concept map tool enables knowledge structurization and visualization.A tool of data processing makes knowledge into understanding data, chart, formula and text.The OLC environment with search tool "allow online learners to determine the browsing sequence, to add to the

information for making it more personal, or to build and structure nodes and links, thereby forming a network of ideas in the knowledge base" (Jonassen, 2000).

Externalization of individual knowledge. Externalization means the learner converts tacit knowledge into explicit knowledge and makes contribution of knowledge to OLC. Externalization is mainly triggered by a social situation, more typical one begins with discussing question or putting forward misconcept.In this way, the learner can put forward new analysis and solution around the problem, externalizes one's own knowledge and explains own viewpoints around misconcept.Generally speaking, externalization is also the reorganized process of individual knowledge. Externalizing individual knowledge also occurs through establishing learning notes, reflective logs, operation programs, visual presentations, E-mails, multimedia, reports and discuss records. Tacit knowledge that the learner externalized is stored in knowledgebase at the same time in order to share and search for other learners.

Socialization of individual knowledge. Socialization of individual knowledge focuses on individual converting tacit knowledge into OLC's tacit knowledge or explicit knowledge, "reinforce shared understanding across the group" (Consway and Whittingham, 2001). In OLC, Knowledge can be tacit (Sorensen and Lundh-Snis, 2001) and become explicit through interaction (Schwen, Kalman et al., 1998), and be transferred through participation in social groups (Sorensen and Lundh-Snis, 2001) too. Learner's experiences and know-how knowledge can be exchanged and shared synchronously or asynchronously through such as conversation, dialogue and meeting in OLC, which strengthens cognition of social sharing and new insight, and so creates and exchanges tacit knowledge. For example, discussion boards help other learners to learn the course topics, integrates knowledge into the learning environments by the common understanding, sharing values, beliefs, languages, and ways of doing things (Trentin, 2001) that forms the basis for discussion and knowledge exchange (Consway and Whittingham, 2001).

4.2 Collaborative Knowledge Innovation Layer in OLC

Knowledge sharing. The concept framework of collaborative knowledge innovation begins with the sharing of individual knowledge. The individual learner enters OLC only through sharing externalized,socialized, combined and external knowledge, and the OLC environment supports learners to create knowledge through exploring other learners' knowledge. Knowledge sharing is a process both which expands individual and whole OLC knowledge storage through knowledge exchanges and which understands knowledge produced in other learner's learning process. Knowledge sharing reflects OLC learners contribute individual tacit and explicit knowledge to OLC, so that other learners can get, thus forming the foundation of knowledge sharing. In OLC, each learner is not only a producer of knowledge but also a sharer of knowledge. Because there are none but all learners' contributed knowledge positively, OLC will have more knowledge

that will be accumulated, thus each learner could share more knowledge. Butler(2001) noted that "the knowledge sharing activity is an important construct for explaining the dynamics of a virtual community since without some forms of community knowledge sharing activity, any virtual community will fail to survive". Reinforcing OLC will offer the ideal chance for the learner to put into knowledge sharing, to keep OLC knowledge being in the state of activating, to develop tacit and explicit knowledge so as to transmit other learners. Cummings (2003) identifies five primary contexts that can affect such successful knowledge sharing implementations, including: the relationship between the source and the recipient,the form and location of the knowledge, the recipient's learning predisposition, the source's knowledge sharing capability, and the broader environment in which the sharing occurs. Generally, there are several ways of sharing knowledge as following:(1)utilizing the forum, chat-room, electronic meeting, E-mail to exchange;(2)question-leading tacit knowledge sharing, namely OLC learners share tacit knowledge on the particular issue and artifacts or communicate their thought processes; and (3) Sharing the produced new opinion after learning from course content and knowledgebase.

Knowledge comparison. Comparison of knowledge needs to examine extant knowledge and knowledge shared, because there has knowledge being identified, valid and reusable in knowledgebase. To the knowledge shared, OLC learners will compare, clarify and ponder over their own understanding and methods of treating knowledge again. Through comparing, learners know about their differences between views and explanations, elaborate questions, clarify concepts, and generate creative collisions which will be the beginning of process of knowledge innovation. Individual learner explains the result of knowledge comparison in detail and posts it on discuss board, and other learners in OLC, through reading,elaborating, questioning and criticizing knowledge,make decisions,appraisals and criticisms of the thought,the fact and the solution.

Knowledge negotiation. Knowledge negotiation focuses on knowledge artifacts in OLC to be developed into the one other acceptable state. After the comparative stage of knowledge, being different views and conflicts, OLC learners engage in the process of negotiation of knowledge aiming at obtaining new knowledge, and they discern difference and similarity, produce abundant, common understanding, and reduce the field disagreed with. Negotiation enables one another more understanding the viewpoint and the question each holds, adopts meta-cognitive statement presenting construction of new knowledge and reflecting fields that are agreed or disagreed with, and reaches the sharable understanding and the common vision. There are three kinds of methods of reaching consensus in negotiation of knowledge: quick consensus building, integration-oriented consensus building and conflict-oriented consensus building (Weinberger and Fischer, 2006). Quick consensus building is a method which the learner can accept others' contribution, not because he/she is convinced or indicates a real change of perspective, but in order to negotiate continuously. Integration-oriented consensus building characterized by receiving the views of

various fields, occurs when the learner gives up or revises initial beliefs and corrects one's own view based on other learners' contributions. An indication for integration-oriented consensus building is that"participants show a willingness to actively revise or change their own views in response to persuasive arguments"(Keefer, Zeitz and Resnick, 2000). Conflict-oriented consensus building is that learners need to more closely operate on the reasoning of others instead of simple acceptance of their contributions, and need to pinpoint out specific aspects of other learners' contributions and modify them or present alternatives.

Knowledge creation. Knowledge creation is a continuous, self-transcending process through which one transcends the boundary of the old self into a new self by acquiring a new context, a new view of the world, and new knowledge(Nonaka, Toyama and Konno, 2000). Effective knowledge creation needs OLC learners to:

- change their cognitive frameworks;
- examine, verify and negotiate innovative knowledge by continuous interaction;
- compare and contrast with the views stated before;
- deal with and synthesize through arguing, consulting and reaching identical knowledge;
- offer views, ideas not considered before or put forward new understanding to existing knowledge;
- discuss the value of view, hypothesis and possible solution;
- strengthen knowledge;
- pay attention to wrong logic and expanded debate;
- promote problem-solving of different situations on the basis of the original, flexibility and deduction;
- produce creative problem solution or external structure of new knowledge.

Knowledge integration. Knowledge integration refers to integrating innovative knowledge or solution of consensus into OLC knowledge that any learner can access, integrating knowledge produced in previous stages to form new knowledge, reconstructing existing thoughts or views from a new perspective, engaging in and reflecting different thoughts that others put forward, connecting their existing knowledge to a series of new knowledge obtained, and building new meaning. Knowledge integration itself is a source of new knowledge which can produce combined different kinds of knowledge resources. The methods of knowledge integration include synthesis, reflection and diffusion, the focus of integration is to make the cooperative effect of knowledge resource promote the innovation ability of OLC, and form the new foundation of application.

4.3 Intermediary Layer of Knowledge Innovation in OLC

Social interaction. Social interaction in OLC includes two dimensions of knowledge innovation and social development. From the perspective of knowledge innovation, personal knowledge and ability are not isolatively but interactively obtained from others who formed the social network. In that context,

knowledge innovation is the creation of knowledge as a social product (Scardamalia and Bereiter, 1996), is a social process and is not merely limited to individual(Nonaka, Takeuchi and Konno, 2000). Social interaction offers opportunities for OLC learner to criticize, prove and help the individual learner to reach what themselves can not, and it enables "covert abstract processes visible, public and manipulable and serves as a necessary catalyst for reflective meta-cognitive activity"(Puntambekar et al.,1997). Knowledge innovation activities are recursive processes including building knowledge, identifying and solving important problems, sharing results, discussing thoughts and making elaboration in OLC. The learner acquires particular knowledge and enhances meta-cognitive abilities of subject through detailing, constructing, collaborating and reflecting. From the perspective of social development, success of OLC is not so reliant on the static 'stock' of knowledge, but rather on the dynamic social processes through which knowledge is enhanced and renewed(Gray and Densten, 2005). Social interaction play an important role in determining that the OLC forms and develops, "social interaction and consequently the social (psychological) processes may give rise to a social space through affiliation, impression formation, and interpersonal attraction that may end in social relationships and group cohesion"(Kreijns and Kirschner, 2001).

Knowledgebase. Knowledgebase as OLC learners' public knowledge assets is the results that OLC knowledge accumulates gradually. Sources of knowledgebase include: knowledge which the teacher prefabricates; knowledge which the learner produces in the course of learning; knowledge which the learner obtains from the outside; and knowledge which the community innovates. It supports knowledge of comparison, creation, combination and decision, solves problem, promotes learners' conversations, and produces potentials such as innovative knowledge. The learner can:

- establish and operate knowledgebase;
- engage in the activity of knowledge expression;
- organize information in one's own way that can be understood;
- promote high order thinking and the meaningful learning connection;
- create new knowledge and form new cognitive structure and schema.

So, the environment of knowledgebase needs learners to reflect personal knowledge, state learning intentions and release thoughts to public knowledgebase. Rights of accessing public knowledgebase are equal to all learners, but the difference depends on how learners interact with knowledgebase and use methods of searching the knowledge.

Cognition connection. The knowledge innovation is a complicated information processing activity which needs learner's cognition participation and cognition connection. The cognition connection refers to cognitive input/output. In order to solve problem in learning and innovate knowledge, the learner needs cognitive activities such as building question space, concept space and relationship between them. OLC culture forces learners to exchange their cognition

connections in an explicit way. The cognition connection reflects the concrete processing activity, and learners must deal with selection, organization and synthesis. Each learner's cognition connection is organized so as to be got by other learners in OLC easily. Generally speaking, there are three kinds of basic activities of cognition connection (Schellens and Valcke, 2005):(1) Presentation of new information. Learners present information that is new in the context of the discussion. Further distinction is made between the presentations of information about facts, experiences or opinions and theoretical ideas. (2)Explicitation. This is a type of communication that reflects a further refining and/or elaboration of earlier ideas. (3)Evaluation. This type of written messages corresponds to a critical discussion of earlier information or ideas. It goes beyond a simple confirmation or negation and reflects argumentations, reasonings, justifications.

Zone of proximal development. The zone of proximal development (ZPD) is the distance between learner's present actual development level and potential development level. In OLC, there are individual ZPD and collective ZPD. In individual ZPD, as to the learner who wants to participate but can't independently create knowledge, he/she needs capable learners in OLC to offer helps or supporting resources, thus creating knowledge which reaches the potential level. In collective ZPD, the learners in OLC can form the intelligence collective, all learners participate and collaborate together to create the new knowledge which surmounts collective ZPD of the whole community by utilizing mutual potentials.

5 Conclusions

OLC is the important environment of knowledge innovation which can realize production, sharing, application and innovation of knowledge. This knowledge innovation is a concept of daily perspective and a simple innovative activity, so there is a great difference from the original activity of knowledge innovation in scientific research.

The model of knowledge innovation in OLC must integrate viewpoints of knowledge creation and management, cognitive and social constructivism. Knowledge innovation in OLC includes two sub-processes of individual knowledge innovation and collaborative knowledge innovation, which are mediated by cognition connection, ZPD, social interaction and knowledgebase. Individual knowledge innovation, which is different from SECI, begins with internalization, via combination, externalization and socialization. In collaborative knowledge innovation, learners of OLC share, compare, negotiate, create and integrate knowledge together.

Acknowledgments. This study was part of the research project of "Models and Practice of Knowledge Innovation in Online Learning Community" by the grant from the 11th 5-Year Plan for Tianjin Educational Development (C023).

References

1. Afhuah, A.: Redefining firm boundaries in the face of the Internet: Are firms really shrinking? Academy of Management Review 28, 34–53 (2003)
2. Butler, B.S.: Membership Size, Community Activity, and Sustainability: A Resource-based Model of Online Social Structures. Information Systems Research 12, 346–362 (2001)
3. Consway, B., Whittingham, V.: Managing Knowledge and Learning at Unipart. Knowledge Management Review 4, 14–17 (2001)
4. Cummings, J.: Knowledge Sharing: A Review of a Literature (2003), http://lnweb18.worldbank.org/oed/knowledge_eval_literature_review.pdf
5. Fahey, L., Prusak, L.: The Eleven Deadliest Sins of Knowledge Management. California Management Review 40, 265–276 (1998)
6. Gray, J.H., Densten, I.L.: Towards an Integrative Model of Organizational Culture and Knowledge Management. International Journal of Organizational Behavior 9, 594–603 (2005)
7. Jonassen, D.H.: Computers as Mind Tools for Schools: Engaging Critical Thinking, 2nd edn. Prentice Hall, New Jersey (2000)
8. Kearsley, G., Lynch, W.: Structural Issues in Distance Education. Journal of Education for Business 71, 191–196 (1996)
9. Keefer, M.W., Zeitz, C.M., Resnick, L.B.: Judging the Quality of Peer-led Student Dialogues. Cognition and Instruction 18, 53–81 (2000)
10. Kreijns, K., Kirschner, P.A.: The Social Affordances of Computer Supported Collaborative Learning Environments. In: 31th ASEE/IEEE Frontiers in Education Conference, Reno, Nevada, USA (2001)
11. Nonaka, I., Toyama, R., Konno, N.: SECI, Ba and Leadership: a Unified Model of Dynamic Knowledge Creation. Long Range Planning 33, 5–34 (2000)
12. Puntambekar, S., Nagel, K., Hbsher, R., et al.: Intra-group and Intergroup: An Exploration of Learning with Complementary Collaboration Tools. In: Hall, R., Miyake, N., Enyedy, N. (eds.) Proceedings of Computer-supported Collaborative Learning, Toronto, Canada, pp. 207–214 (1997)
13. Scardamalia, M., Bereiter, C.: Student Communities for the Advancement of Knowledge. Communications of the ACM 39, 36–37 (1996)
14. Schellens, T., Valcke, M.: Collaborative learning in asynchronous discussion groups: What about the impact on cognitive processing? Computers in Human Behavior 21, 957–975 (2005)
15. Schwen, T.M., Kalman, H.K., Hara, N., et al.: Potential Knowledge Management Contributions to Human Performance Technology Research and Practice. Educational Technology Research and Development 46, 73–89 (1998)
16. Seufert, S.: Design and Management of Online Learning Communities (2002), http://www.aib.ws.tum.de/euram/seufert_paper.pdf
17. Sorensen, C., Lundh-Snis, U.: Innovation through Knowledge Codification. Journal of Information Technology 16, 83–97 (2001)
18. Starbuck, W.H.: Learning by Knowledge Intensive Firms. Journal of Management Studies 29, 713–740 (1992)
19. Trentin, G.: From Formal Training to Communities of Practice via Network based Learning. Educational Technology 41, 5–14 (2001)
20. Weinberger, A., Fischer, F.: A Framework to Analyze Argumentative Knowledge Construction in Computer-Supported Collaborative Learning. Computers & Education 46, 71–95 (2006)

The Design of Software Architecture for E-Learning Platforms

Dongdai Zhou[1, 2, 3], Zhuo Zhang[1], Shaochun Zhong[1, 2, 3], and Pan Xie[1]

[1] School of software, Northeast Normal University, China, 130024
[2] Engineering Research Center of E-Learning Technologies, Ministry of Education, China, 130024
[3] E-learning laboratory of Jilin Province, JiLin ChangChun, 130024
ddzhou@nenu.edu.cn,sczhong@sina.com

Abstract. Although e-learning have been widely used at schools, universities and other institutes, some obvious shortcomings have been recognized. Current e-learning platforms are developed using existing technologies and compensate for the disadvantages of traditional education methods. However, most of them are not flexible and efficient enough to support real world teaching and learning. In this paper, we present flexible and hierarchical reusable dynamical software architecture for e-learning platforms and an approach of software components integration based web services. We also analyze how this architecture can facilitate a web-based e-learning system development. **Keywords:** E-Learning; Software Architecture; Software Product Line; Web Services

1 Introduction

E-Learning is just-in-time education integrated with high velocity value chains[7]. It is the delivery of individualized, comprehensive, dynamic learning content in real time, aiding the development of knowledge communities, linking learners and practitioners with experts. Generally, e-learning improves the flexibility and quality of the education by[6] :

- providing access to a range of multimedia resources, such as graphics, sounds, animations and videos;
- supporting the reuse of high quality and expensive resources;
- supporting increased communications between instructors and students and between students;
- enabling instructors to provide different materials to the students from different backgrounds;
- encouraging students to choose materials according to their own interests and to study at their own pace;
- encouraging students to take responsibilities for their own studies.

So far, a number of research teams have implemented different kinds of Web-based education platforms to support learner-centered, interactive and active learning. The Web is used not only as a delivery medium but also to foster free exploration of

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 32–40, 2008.

learning materials and to allow the learners to interact with materials, instructors and other learners. Almost all the systems fall into three categories: Courseware developing System(such as The Geometers Sketchpad, and Authorware), teaching supporting system(such as Blackboard[1]) and learning supporting system(such as Web CT[2]) , and web-based educational resources system or portal (such as NGFL). Although those e-learning systems provide versatile functions to support e-learning and are widely used by schools, universities. They, however, have the following shortcomings:

- Current e-learning systems are teacher-centered. They are designed to facilitate teaching activities but not consider the learners' diverse learning goals and demands. The learning process is not well-supported.
- Most e-learning systems are not developed for customization. They cannot meet all requirements of instructors and learners in their teaching and learning activities. Meanwhile, most instructors of K-12 are not good at developing software, and cannot build a complicated software system on their own.
- Most e-learning systems are not structured in a perfect architecture, and their elements are tightly coupled. So there are a lot of problems such as redundant development, difficulties to integrate with others and maintenance difficulties among e-learning software products.

Because of these shortcomings, current e-learning systems have limited applications in supporting real world teaching & learning. In our e-learning platform, we are engaged in improving the performance of e-learning through building a flexible and hierarchical reusable dynamical application architecture and present a service-oriented approach of software components integration. Section 2 introduces the architecture design of our Platform. Section 3 specifies the method of integrating different software components. Section 4 specifies an application of our software architecture for e-learning platform to develop a web-based special topic learning website. Finally, the paper is concluded in Section 5.

2 The Software Architecture for E-Learning Platform

This Software Architecture is a flexible and hierarchical reusable architecture based on domain-specific software architecture(DSA) and software product lines(SPL)[8].Under this architecture, we encapsulate business logic into software component firstly, and then weave the unchanged or uneasy changing parts of the e-learning domain logic into respective domain frameworks , and margin the easy changing parts as pluggable user interfaces in order to insert or delete or replace the components according different requirements in the future. Subsequently, we build several e-learning software product line based the domain frameworks. Consequently achieve the goal to fast-building and integration of e-learning systems via those e-learning software product lines. By using this software architecture, instructors can build customized application software systems in a visual studio just like building block.

This Architecture is divided into four levels: the application layer, the product line layer, application framework layer, component library layer. It is shown in Fig. 1.

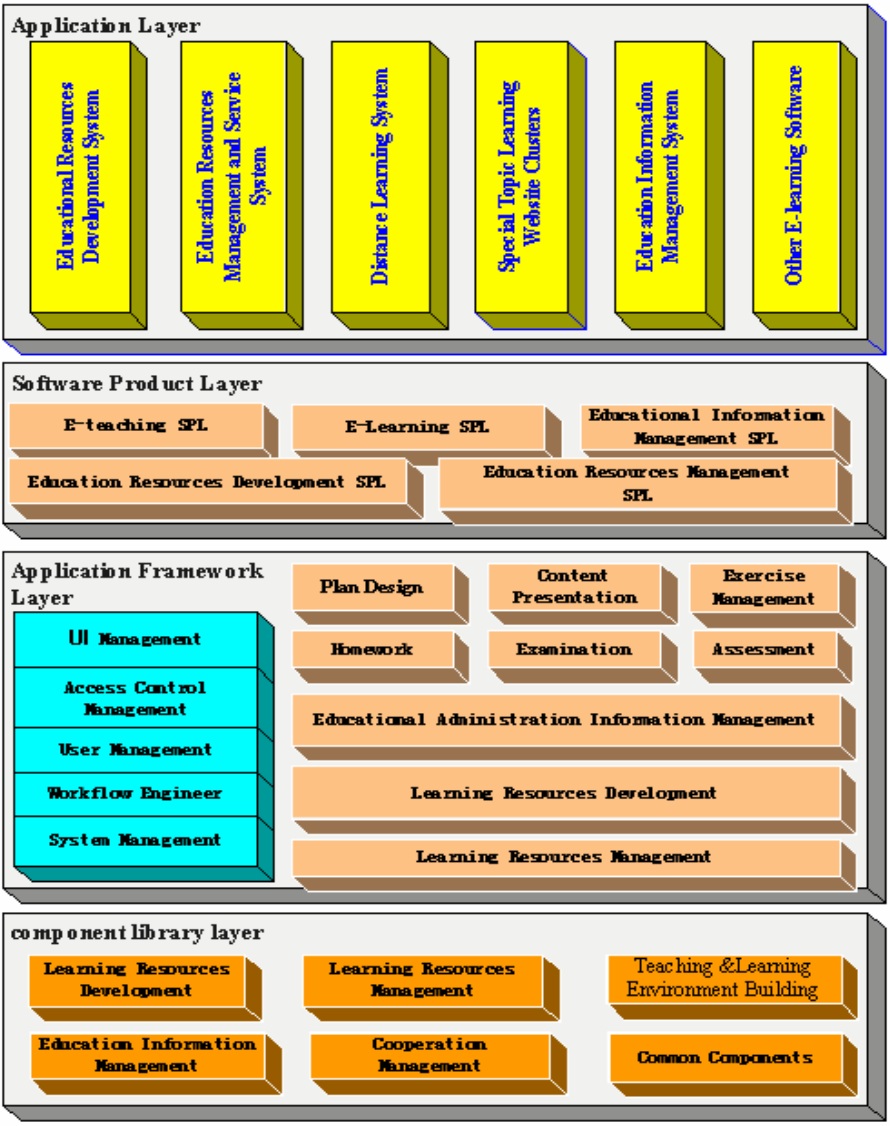


Fig. 1. The E-learning Platform Software Architecture

(1) Product line layer: It mainly includes five SPL, which are e-teaching supporting SPL, e-learning supporting SPL, education information managing SPL, education resource development SPL and education resources management and service SPL. This layer is responsible to provide reusable domain solution of e-learning, describe the functional and non-functional requirements of components which composed the various product line architectures and the associate relation between these components, and realize the building of specific software application system. This layer is the core of

the platform software architecture. Each product line is designed for a specific goal of domain application and can support a group of buildings of similar application systems. Through selecting and assembling different frameworks and components, Instructors can quickly instantiate a customized software system.

(2) Application Layer: This layer is a specific group of e-learning software systems which are built via different product lines.

(3) Application framework layer: The layer is the collection of the public frameworks and domain specific frameworks. Each framework is composed of a series of related components, component associated relations and restrictions. It is the collection of associated components which solute certain sub problem of the product line. The public framework mainly includes UI management, access control management, user management, workflow engineer, system management, all of which are common frameworks to all the product architecture. Meanwhile, the domain specific frameworks are designed according to different teaching&learning activities design. It covers teaching&leaning plan design, and contents presentation, and exercise management, and homework management, and examination management, and question&answer management, and assessment management, and learning resource development, and learning resource management, and educational administration information management. By introducing the application framework layer, we reduce the complexity of the architecture of product line and increase the reusable efficiency greatly.

(4) Component library layer: This layer is the foundation of the platform software architecture. It is a reusable collection of software units which is already validated by other projects. It mainly includes the component library of educational resources development, and the component library of educational resources management, and the component library of building teaching&learning environment, and the component library of educational information management, and the component library of cooperation, and the common component libraries for software system management. Moreover, the component library layer is extensible and distributed. As long as the third-party components meet with the certain agreement of application framework, they can be added to the component library layer.

3 The Approach of Component Integrating and Framework Weaving Based on Web Services and XML Data Bus

In our e-learning platform software architecture, the application framework is the collection of components which solute certain sub problem in a specific software product line. It provides a group of basic constitution units to establish software product family. Prior to this, there are two kinds of framework integration techniques: the traditional object-oriented framework[3] and component-based framework[4]. In the former method, the framework is often referred to the classes in the object-oriented programming, and those classes are integrated into an application framework based on the inheritance and interface mechanism of class. This kind of framework has three main shortcomings: maintenance difficulties, and small granularity of reusable software units, and depending on the specific programming languages. As to the framework based on components, the components are composed units which have

well-definition interface and prescribe dependency relationship in the context clearly. This approach has a list of good characteristic, such as reusable granularity bigger than classes in OO technology, better encapsulation, non-depending on programming languages. However, this approach is tightly coupled. It need define the interface and dependency relationship clearly, further more if the interface of one component changes, all the components connected with it will fail. It makes against the maintenance and evolution of framework.

Therefore no matter we adopt the framework based on object-oriented framework or components to build domain framework in our software architecture, the efficiency of the platform will be decreased. To solve this problem, we present an approach of component integration and framework weaving based on Web Services and XML data bus.

The main principle is based on the standard protocols of UDDI/WSDL/SOAP to encapsulate all kinds of e-learning objects into reusable Web Services components[5]. Meanwhile, we build a XML data bus for the coordination and communication between components to realize components plug-in or bind dynamically. The XML data bus is a data pool of data entities, and is composite of both XML and XML Schema. In this data pool, the XML structures represent an existing set of classes (which deal with business data in an object model), and the XML Schema is used to

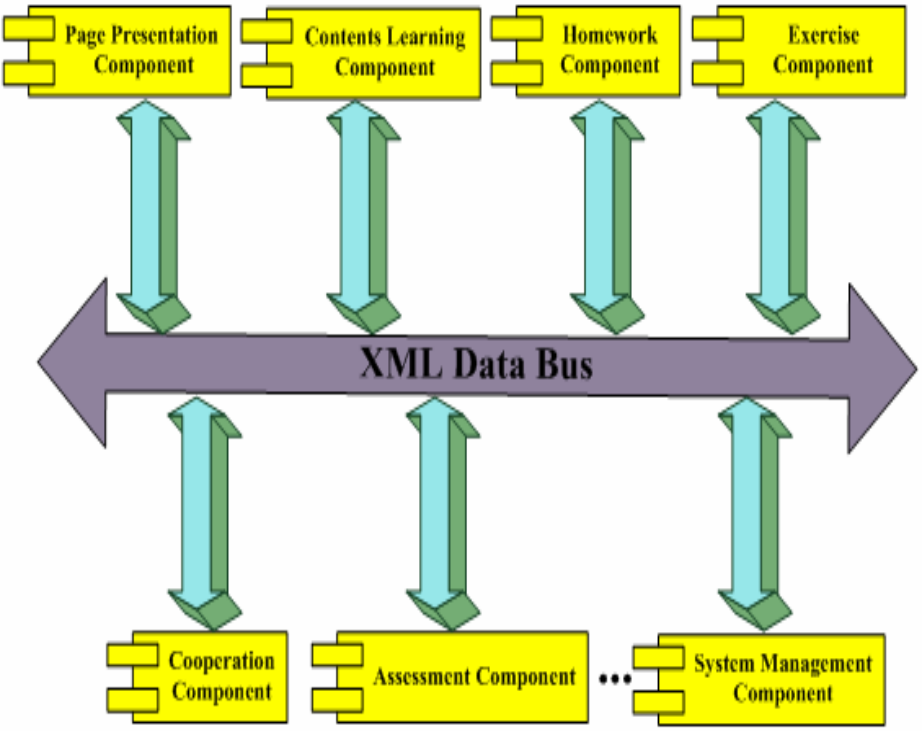


Fig. 2. The model of component integration and framework weaving

build the object models for data persistent from XML structures. Based on XML data bus, each component (encapsulated into web service) only read and write the XML data to accomplish its business logic without the need of communicating with others component. The XML data bus is controlled by a data bus controller, which is in charge of Marshalling and Un-Marshalling. Marshalling refers to the process that takes an object or a tree of objects, and creates the XML representation that will record its state. It's a kind of serialization.Un-marshalling is just the opposite, it takes the XML representation and builds an object model from it for object persistent. Once this data bus is in place, it is easy and quite straightforward to integrate the different domain components and weave them into a black-box framework, which is dynamical and compatible and supporting of thermal plugging. Moreover, on inheriting the characteristics of Web Services, the purpose of loosely coupled and independence of programming languages and operating system also be reached. Thus users can only concentrate on the business logic and all the other issues in their domain project without having to concern themselves with the technical details of application development. Figure 2 shows the principle of component integration and framework weaving based on XML data bus.

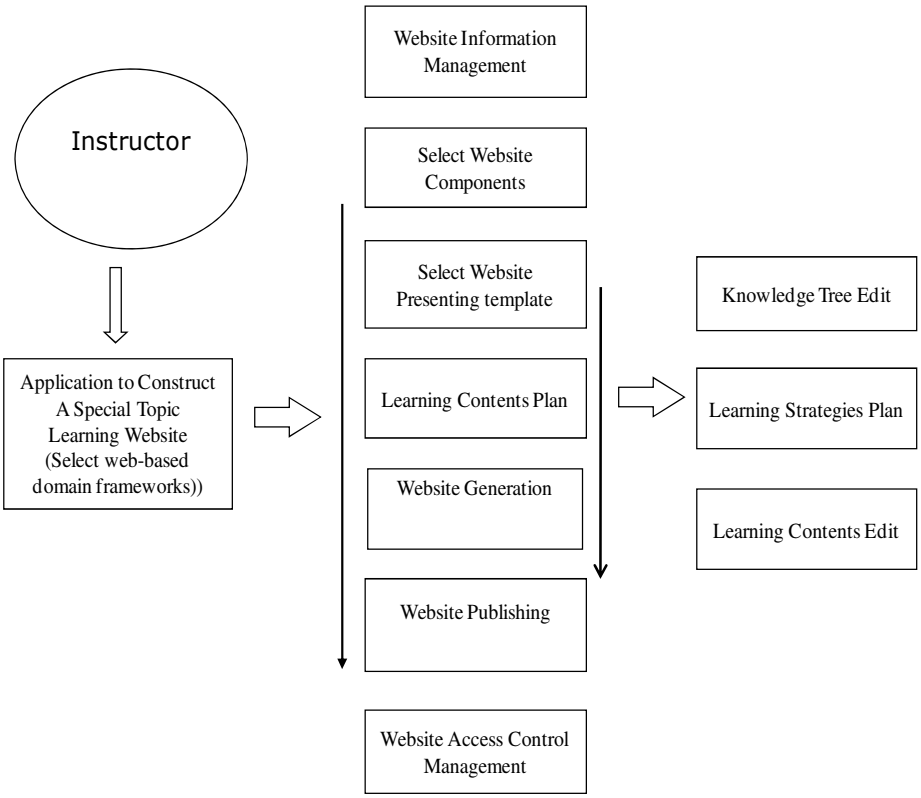


Fig. 3. The Logic Flow to Build a Special Topic Learning Website

4 The Development of a Web-Based Special Topic Learning System

Based on the above e-learning software architecture, we developed a web-based special topic learning website system. Using this topic learning website system, Instructors don't need to learn the knowledge of software design and programming. They can just choose e-learning software product line, and then select web-based domain frameworks and components to build various subject special topic learning websites, which covers learning content management, and homework, and exercise, and test, and question & answer, and memo, and cooperation and so on. Moreover, the website will be generated and published by the system automatically. Figure 3 gives the logic flow to contruct a web-based special topic learning system. Figure 4 gives the building UI of special topic learning website. Figure 4 is the building result of a Success English Conner Topic Learning website.

Fig. 4. User Interface of Building a Special Topic Learning Website

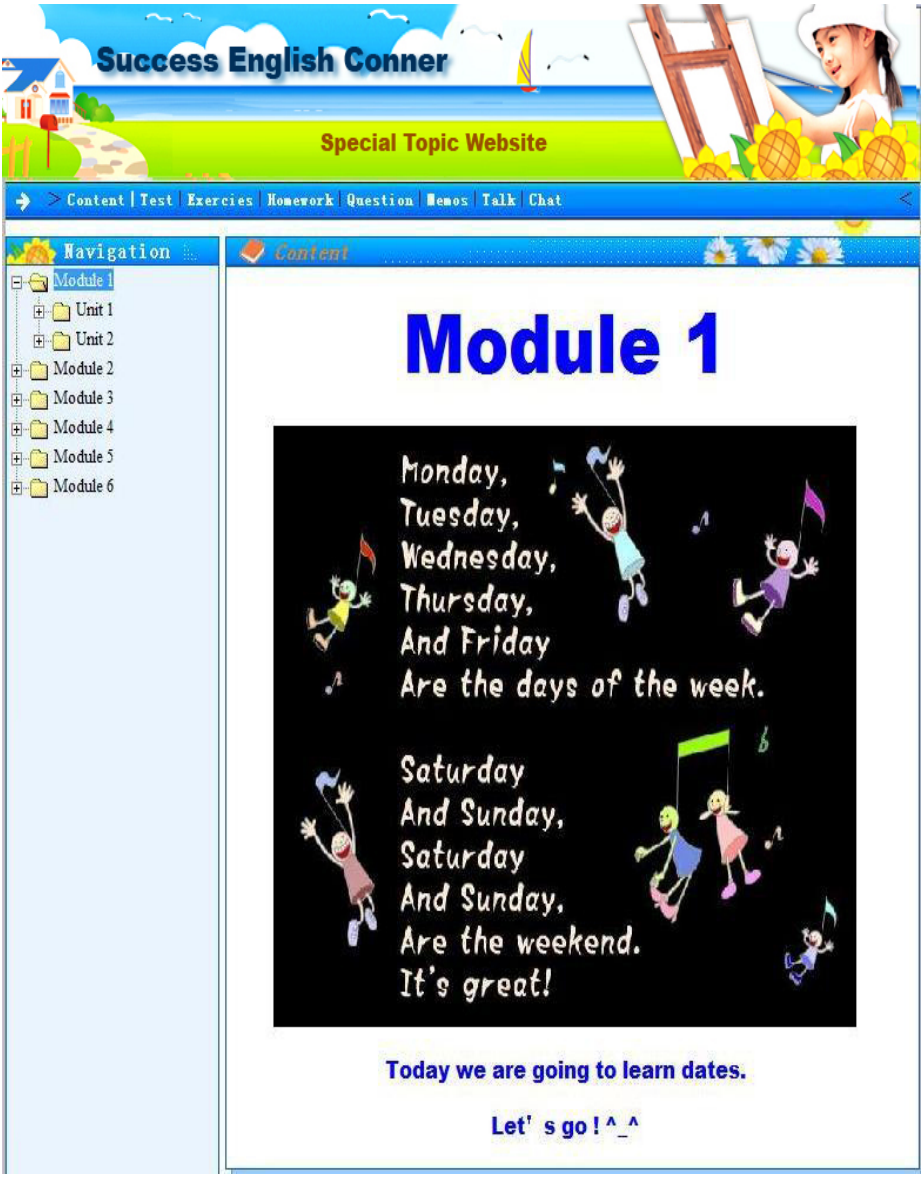


Fig. 5. Learning Contents Presentation Page of Success English Special Topic Learning Website

5 Conclusion

In this paper, flexible and hierarchical reusable software architecture for e-learning platforms is introduced. It consists of four layers: the application layer, the product line architecture layer, application framework layer, and component library layer. These are responsible for building customized e-learning systems just like building block

respectively. In addition, an approach of component integrating and framework weaving based on web services and xml data bus has been outlined. The experiment results demonstrate the satisfying effect of this software architecture.

For future work, the reliability will be improved. More e-learning activities will be supported. It will be easier for instructors to build customized digital learning environment to different students. The methods of evolution of this e-learning software architecture will be researched as well.

References

[1] Blackboard (2005), http://www.blackboard.com/
[2] WebCT (2005), http://www.webct.com/
[3] Mattson, M.: Evolution and Composition of Object-Oriented Frameworks, printed in Sweden, KaserntryckerietAB, Karlskrona (2000)
[4] Wenhui, H., Wen, Z., ShiKun, Z.: Study of Application Framework Meta-Model Based on Component Technology. Journal of Software 15(1) (2004)
[5] Long, W., Zhong, S., Zhou, D.: A Distance Education System Based On Web Services. Journal of Computational Information Systems 2(1), 139–144 (2006)
[6] Zhou, D., Cheng, X., He, X.: The Development of a Customized E- Learning System. Journal of Computational Information Systems 2(1), 211–216 (2006)
[7] Ducker, P.: Need to Know: Integrating e-Learning with High Velocity Value Chains, A Delphi Group White Paper (2000), http://www.delphigroup.com/pubs/whitepapers/20001213-e-learning-wp.pdf
[8] Anastasopoulos, M., Atkinson, C., Muting, D.: A Concrete Method for Developing and Applying Product Line Architectures. In: Aksit, M., Mezini, M., Unland, R. (eds.) NODe 2002. LNCS, vol. 2591, pp. 294–312. Springer, Heidelberg (2003)

An Educational Component-Based Digital TV Middleware for the Brazilian's System

Juliano Rodrigues Costa [1,2] and Vicente Ferreira de Lucena Junior [2]

[1] Genius Institute of Technology, Av. Dr. F. Coelho, 64 São Paulo – SP – Brazil 05423-911
jcosta@genius.org.br
[2] Federal University of Amazonas; Ceteli – Electronics and Information Technology R&D Center - Av. Gen. Rodrigo Otávio, 3000 Manaus – AM – Brazil 69065-190
vicente@ufam.edu.br

Abstract. One of the major problems faced by the Brazilian population is the low level of the fundamental schools. Television is the most popular source of entertainment and information of the Brazilian population being present in approximately 54 million families all over the country. These families watch television for more than 8 hours daily. Moreover, at this moment, the Brazilian TV system is moving from analog to digital. That means not only that image and sound will be delivered with much better quality but also that it will be possible to send interactive multimedia programs, creating a brand new way of watching TV. That is in fact the main novelty of the digital system it will be possible to offer personal interactive services such as banking, games and most importantly educational programs. This work introduces a software framework called "Extended Middleware for Digital TV (EMTV)" which is suitable for the generation of interactive applications executed over digital television systems. Its concept was developed focusing on the Brazilian technological options for Digital TV. Technically, EMTV is a procedural GEM compliant application which, from the programmer's point of view, acts as a declarative middleware extension. The framework was developed to be component-based in order to minimize the need for programming knowledge to deploy the digital TV applications using EMTV. The main goal of the platform is to facilitate the construction of interactive multimedia educational applications, a crucial field for the Brazilian population. The concept is tested and validated by the construction of a Quiz application presented at the end of the paper.

Keywords: Multimedia Interactive Digital TV, Educational Applications, Digital TV Middleware, Component-Based Software Development, Quiz.

1 Overview of the Brazilian DTV System

Brazil is at a point in time where an important technological decision will affect the life of 90% of its 184 million citizens who consider television as one of the most important sources of information and entertainment. This decision refers to the use of digital technology in the current process of transmitting and receiving open TV signals in the country, which was started in December 2007.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 41–51, 2008.

It is only in 1998 that Brazil started to research DTV technology and initially decided to develop its own standard which used to be called SBTV (Brazilian TV System) and whose main characteristic is the use of an OFDM (Orthogonal Frequency-Division Multiplexing) modulator system equipped with an artificial intelligence module on the reception side to make the receiver multipath-noise robust. Despite the good results achieved the a few years ago, the Brazilian government showed clearly that it was giving in to the pressure of the TV content providers when it decided to adopt the Japanese ISDB (Integrated Services Digital Broadcasting) standard. At that time, the Brazilian committee renamed the standard to ISDTV [1] (International System for Digital TV) on account of some local contributions. As a matter of fact this model, already in use in Brazil since last December 2007, consolidates important aspects of the Japanese technology such as the use of the BST-OFDM (Band Side Transmission – Orthogonal Frequency-Division Multiplexing) modulation which is very effective against multipath noise even with fast-moving mobile receivers[1]. The Brazilian contributions [2] relate to the use of the MPEG-4 AVC standard also known as MPEG-4 "Part 10" or H.264 which is an evolution of the MPEG-2 standard primarily used in ISDB because it achieves compressed audio and video rates 40% to 70%[2] higher than those of the MPEG-2. With such characteristics, and considering that the Japanese standard committee is aggregating the proposed changes, the Brazilian model was recently defined to be referred to internationally as International ISDB.

Another Brazilian contribution to ISDB refers to the development of its DTV middleware specification called GINGA [3]. Just like other DTV middleware it has a procedural part, the GINGA-J (Ginga – Java) [1], and a declarative part known as GINGA-NCL [1] (Ginga – Nested Context Language).

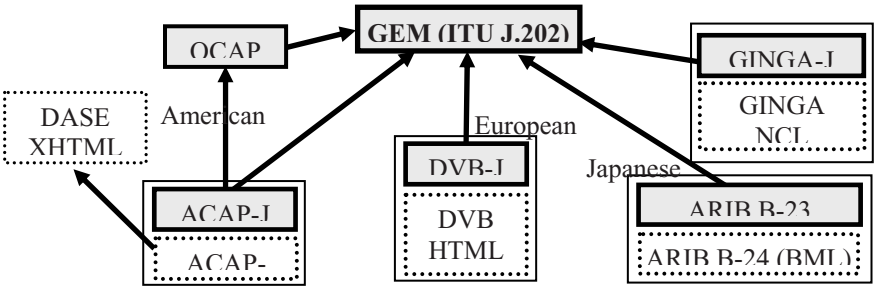


Fig. 1. Relationship between DTV Middleware Specifications

Figure 1 illustrates that the most important middleware specifications available nowadays has a procedural part, represented by the grey boxes, and a declarative part, represented by the white-dashed boxes. It also shows a tendency, the procedural middleware has to be GEM [12] (Globally Executable MHP) compatible. Despite the

[1] HDTV (1920x1080i 16:9) in fixed terrestrial TV receivers. LDTV (320x240 4:3) in fast moving mobile receivers.

[2] Depending on the MPEG-4 profile and on the nature of the images.

fact that GINGA-J has not yet been officially released it will probably be a GEM implementation. GINGA-NCL is the declarative part of GINGA and is considered to be very powerful and flexible as it not only controls appearance and the positioning of media objects but also considers the temporal relationships between them [1].

Terrestrial open DTV transmissions started in Brazil without any middleware support. According to the local industry, this was due to the need to the middleware to support the H.264 standard and due to the wait for official GINGA specifications. Anyway, even without the support to any type of middleware, the innovations used by the Brazilian system made both the signal generation equipment and the set-up boxes very expensive, a fact likely to be a problem for the popularization of the technology in the country.

Nevertheless, the new interactive features of the digital system are expected to have great impacts on the population, the most significant being the use of this technology as a tool to contribute to educational processes. This paper introduces a component-based framework, named "Extended Middleware for Digital TV (EMTV)", which was developed at a time when the GINGA was not even available for public download and whose main target is to help TV-content providers with no advanced programming skills to deploy DTV applications for educational purposes more specifically. The EMTV is free for any use and does not require any expensive tool since it thereby hopes to contribute in democratizing DTV technology in Brazil.

2 Used Concepts on the Extended Middleware

The development of complex DTV applications requiring the procedural approach is a relatively difficult task. Besides the logical concepts, the programmer must have borad knowledge of several software interfaces and must be able to build a very efficient software code due to the hardware memory and processing limitations [4]. The programmer also has to predict all necessary software responses to any user actions and system errors. A DTV application cannot, in any circumstances, force the user to reset his/her television set. This is why, although the procedural approach is powerful, it demands the professional services of an experienced programmer with software engineering capabilities.

The development of applications using declarative middleware [5], on the other hand, is simpler than the procedural applications, since their main functionalities are internally programmed. And so, programmers won't have to concern with most of exception handling once it's already treated by the middleware itself. The limited number of functionalities has the advantage of guarantying the simplicity despite of being less powerful. Declarative middleware only becomes more complex as the offered number of functionalities and flexibility increase. That is a problem most declarative middleware have nowadays.

Figure 2 shows that EMTV has an extra software layer to run over the available middleware in the STB, thereby extending its abilities and allowing it to be able to generate a specific family of DTV applications. In addition, EMTV is designed to be user-friendly, as the knowledge required on software libraries and the executor system is reduced to a minimum. In terms of applications, this additional software layer assumes the middleware attributions once it keeps converting data information into

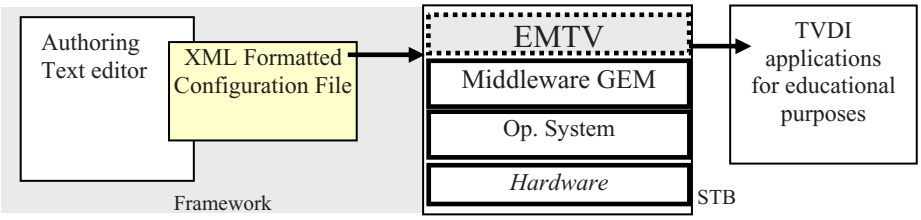


Fig. 2. EMTV framework

interactive applications. The EMTV framework uses an external file which format is defined in the framework. The handling of this file demands the use of any of the third-part text editors referred to as authoring editors in Figure 2.

2.1 Declarative Approach

To minimize the need of high qualified specialists on DTV software development, allowing program providers, more concerned with the content to be presented than to software programming, EMTV offers a declarative approach. It follows the same idea as GINGA-NCL but, is much more simplified and targets a very specific kind of application so that EMTV is able to provide just enough flexibility to deal with the main demands of a specific type of DTV application.

Although the programmer will view the platform proposal as declarative, it was developed using a GEM compatible procedural middleware. This choice is very convenient not only because GEM offers all the necessary library software [12] to develop any desirable feature but also because it allows EMTV to run over any compatible middleware, including MHP, ACAP, ARIB B-23 and especially the Brazilian GINGA-J which is the purpose of this work.

The decision to develop EMTV through the resources offered by GEM specification defines that the platform is in fact a Java XLet [13], meaning that its execution processes can be done just like any other GEM compatible application: The application will first be transported via the data channel or via the return channel[3] before being loaded and managed by the GEM middleware and controlled by DTV or user events [7]. At this point, the EMTV assumes its task as a middleware extension, loading text and image content.

2.2 Educational Purposes

Interactive applications have already been used in several countries where DTV technology has been implemented. As the range of applications increases, new terms like T-commerce, T-government and T-learning and others have been created to classify it according to its main purpose. The term “T-learning” [6] refers to the use of DTV technology in educational processes, one of the main purposes of EMTV. T-learning is viewed as the convergence between DTV technology and E-learning

[3] Starting from MHP 1.0.3 version the XLet applications can be loaded directly from the return channel connected to a TCP/IP network if available.

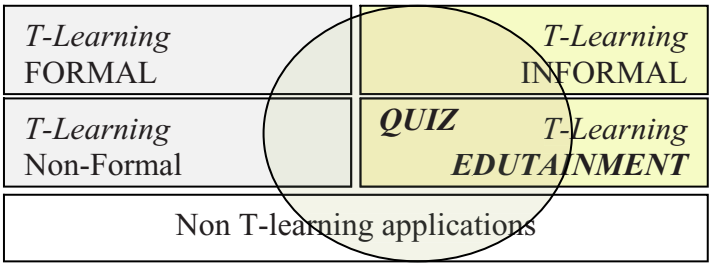


Fig. 3. Proportionality of Quiz Applications in the Universe of T-learning Applications

which basically refers to the use of computational technology for training or any other educational activity. It comes from "E-learning" the popularization of Quizzes as a very effective educational tool.

Walldén and Soronen [10] wrote a paper about relationship between e-learning and t-learning which classifies education processes in four classes: Formal learning (leads to acknowledge diplomas), Non-formal learning (education from formal institutions but that does not provide official diplomas), Informal learning (education which comes from social activities) and Accidental learning also known as edutainment (education + entertainment: the knowledge or skill is acquired not intentionally). A Quiz is a kind of test containing a series of questions along with some alternatives which the users can select according to their skills, knowledge or personal opinions. This principle allows Quizzes to be used not only in any educational process, typically in edutainment and informal processes, but also in other non T-Learning applications.

2.3 Designed to Be Configurable and Easy to Reuse

The need of an external file is already part of the declarative approach. It allows EMTV to be easily reused, as different configuration files generate different applications with a different content and behaviour. The EMTV configuration file is based on the XML format which furthers three main advantages at least:

- XML is easily read and written by both humans and algorithms.
- XML can be validated through DTD files.
- Several free XML editors can be used for edition and validation.

Another advantage of the use of the XML format for the configuration file is that it is very convenient to describe the software components properties. Software components are artifacts constituted by one or more instances of classes, independent enough to provide some interest functionality such as visual, behavioural or both. The behaviour and characteristics of a software component are defined by the properties it makes available through public methods which constitute the interface of the component. There are many good software-engineering proven reasons [11] to design components-based software. EMTV is component-based mainly because it contributes to an easy system reuse, maintainance and expansion.

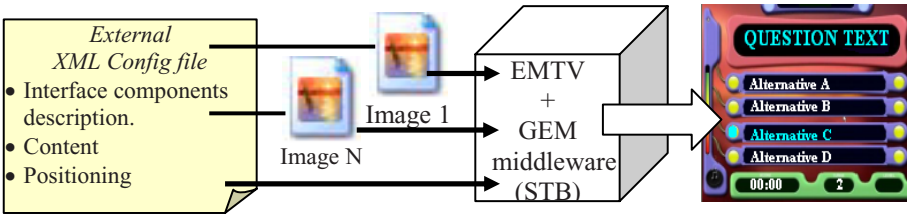


Fig. 4. The EMTV configuration files provides enough information to generate the applications

3 EMTV Components

The first version of EMTV provides only the minimal set of components to allow the easy deployment of TVD Quiz applications with the possibility to use or not a return channel. Those components were defined from a study of similar systems, such as MOODLE [8] which has a module to generate Quiz applications typically for the Internet.

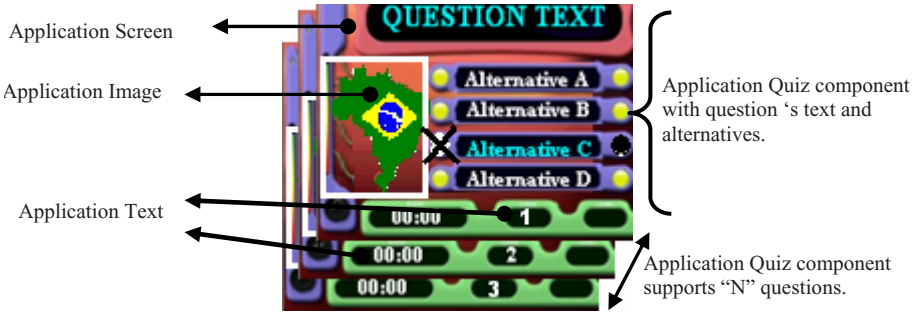


Fig. 5. Quiz application interface that helps to identify graphical components for EMTV

Figure 5 helps to identify some of the basic components to build a Quiz application. It leads to construct 5 basic visual components identified as: Application Screen, Application Text, Application Image, Application Image Button and Application Quiz. There is also a non visual component called "Application Communication" to perform communications between EMTV and an external server. The user indicates the components in the XML configuration file which is read[4] by an EMTV Application Manager class that creates and controls the components according to the configuration file.

3.1 Application Screen

It is a singleton graphical component to represent the screen on which all other graphical components are placed. It is built using HAVI, DVB and JMF libraries and

[4] EMTV uses nanoXML library to be able to read the XML file attributes. This library is a very small and efficient XML parser for Java and can be found at http://nanoxml.cyberelf.be/.

allows the programmer to control the size, position and appearance of the application. One of its attributes indicates if the screen background should be filled with an external picture, giving to the application an enhanced appearance.

3.2 Application Text

It is a graphical component to represent texts over the screen. It is made up by HAVI and DVB libraries but mainly by the instantiation of the org.havi.ui.HText HAVI Class. The attributes allow the programmer to control the text, position, font, font size, foreground colour, background colour and also has a field to indicate some basic animations as blinking and scrolling. Another field also allows the programmer to use some keywords to build a Boolean expression to controls the visibility of the component. The same principle is applied in the Text-content field as EMTV interprets some keywords which are related to information of the running application like the date and time, information about Questions Group application navigation and about the sending status of an Application Communication component. The programmer can place as many Application Text components as needed.

3.3 Application Image

It is a graphical component that is used to represent static images over the screen. It is made up by HAVI, DVB and JMF libraries but mainly by the instantiation of the Java Image Class. The attributes allow the programmer to indicate an external image and controls its position on the screen. The loading method of the image is done by JMF through the classes java.awt.MediaTracker and java.awt.Toolkit. Just like the Application Text, there is a field to control the visibility of the component. The programmer can place as many Application-Image components as needed.

3.4 Application Image Button

This is just a specialization of the Application Image component. It quickly switches between two defined images when the user presses on a key that is also defined on the remote control. This component is just useful to provide a graphical effect like a button if the proper pair of images is used.

3.5 Application Questions Group

This is a singleton component which enables EMTV to build Quiz applications. Its fields provide all the information to create a multi-page Quiz. Each page contains one question, represented by an HText Havi instance, and multiple alternatives for this question. The programmer can define the position, the foreground colour, the background colour of the question whose text is updated every time the user changes the question page through the remote control left and right arrows. The programmer can indicate a different number of alternatives for each question and if each question has a single or multiple choices answer. Fields control the initial position for the first alternative of all questions as well as the relative increment on the x and y axis for the

next available alternatives. Other fields indicate the foreground colour for the alternatives, the foreground colour used as the user navigates through the alternatives, using the remote control up and down arrows, as well as a field to indicate an external picture which will be placed close to the chosen alternatives by pressing on the ENTER key of the remote control. Each alternative has special attributes which allows the programmer to define if the alternative is in fact an edit box where the user can input alpha-numeric, numeric or password characters when the alternative, by default represented by an org.havi.ui.HText is replaced by an org.havi.ui. HSinglelineEntry instance. This last situation is useful for the Quiz to get specific user information, like the username to be registered on an external server if an Application Communication component is used. In this case, an alternative can be converted into a text instruction for the edit box if the selectable attribute for an alternative is indicated as false. The component always inserts an additional configurable question which allows the user to indicate he/she has given the answers. This event disable the possibility of user change its responses and also starts the sending of the collected information in case of the presence of an Application Communication component. After this event the component changes the background color of all alternatives marked with the "isAnswer" attribute allowing the user to check his score.

3.6 Application Communication

It is a non-graphical singleton component which tries to establish a TCP/IP communication through a permanent or dial-up interface, if available, to send the information to an external TCP/IP server. Its fields provide all necessary information to connect EMTV to an external server. The data field replaces several reserved keywords into application specific information including the answers captured in an Application Question Group component and other useful pieces of information. This component is built based on org.davic.resources, java.net.Socket, java.io.DataOutputStream and java.io.DataInputStream classes.

Figure 6 shows the main classes of the EMTV middleware.

4 EMTV Tests and Validation

EMTV was tested and validated at the Ceteli DTV laboratory of UFAM as shown in Figure 7 and 8. The process used an external Apache server connected to the same TCP/IP network as an interactive profile MHP Set-top box (STB). The server is used to provide the EMTV packet application as requested by the STB. Once loaded in the STB, the EMTV requests the server the configuration file and all necessary picture files to generate the Quiz to be presented on the HDTV monitor. The user interacts with the application though the STB remote control and once he/she has finished, the information gathered is sent to an external Apache/PHP server. In the case of this test, the server used to provide the application was also used to register the users' responses.

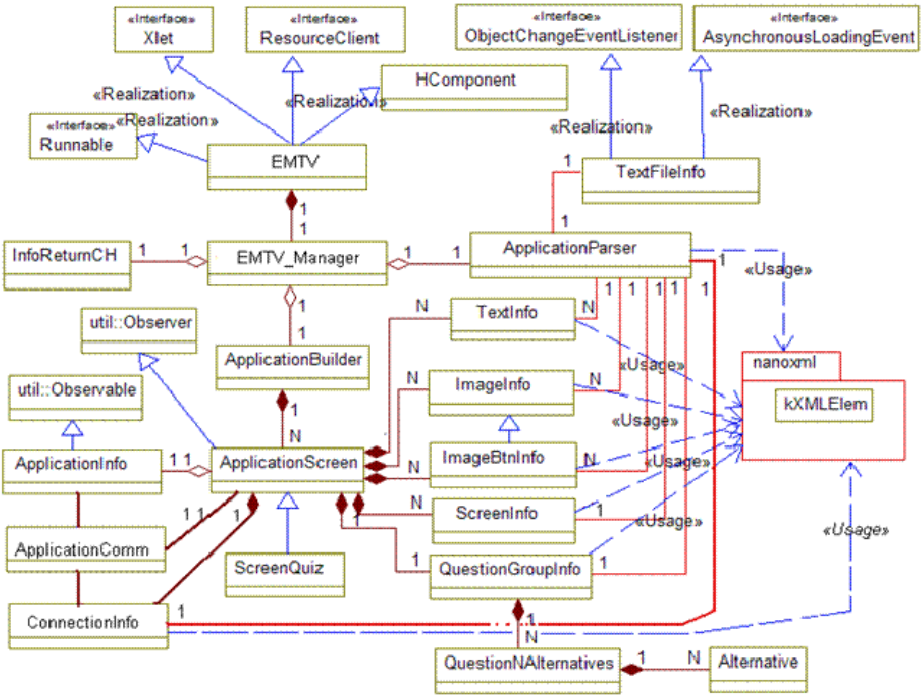


Fig. 6. Class Diagram of the main Components of the EMTV

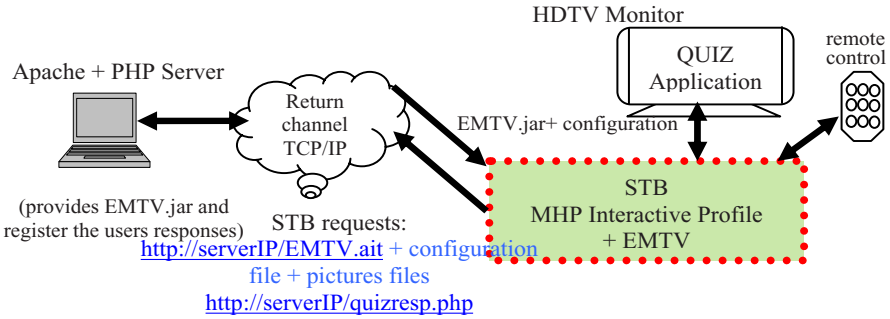


Fig. 7. EMTV tests schema

The same results could be achieved loading the EMTV application directly from the DTV data channel. To validate EMTV on this case would be necessary, besides the proper hardware equipment to generate the MPEG-2 Transport Stream signal and the DSMCC (Digital Storage Media Command and Control) system, to correctly setup the PMT (Program Map Table) and the PAT (Program Association Table) so the STB can detect and load EMTV. This test would also require a minimal change on the software so EMTV would be able to download any external files through the DSMCC synchronously or asynchronously.

Fig. 8. Pictures of a Quiz application generated by EMTV running in a real DTV system

5 Conclusion

Since last December the Brazilian TV broadcasters start to transmit their signals in digital form. Many good things have been written about this fact and the local population is anxious for taking advantage of this new technological facet of a very popular entertainment vehicle. More than a better image and sound it is expected to make good use of the interactive possibilities available with the new digital system.

Knowing the big deficiency in education all over the country and the cultural relationship with TV we can assure that the new digital TV system will certainly play a very important role in the generations to come. Certainly there will be lots of multimedia interactive applications available for usage in a near future. That is exactly the main contribution of the EMTV middleware for interactive digital TV systems presented in this work.

EMTV is a very simple to use and already capable to generate a series of useful multimedia applications for educational purposes. It has the advantage of neither requiring any broad knowledge on DTV systems, nor the purchase of expensive proprietary tools, and thereby encourages the popularization of a generation and use of interactive DTV software.

EMTV is available for free, for any use, by sending an e-mail request to the authors. The following steps of this work will include the development of a graphical authoring tool, as well as the development and improvement of new graphical components.

Acknowledgements

We would like to thank you very much the students involved in the development of this work. Many thanks for the whole CETELI/UFAM staff for giving the laboratorial support.

References

1. Soares, L.F.G., Souza, G.L.: Interactive Television in Brazil: System Software and the Digital Divide. In: EuroiTV (2007)
2. Souza, G.L.: Standard 06 - ISDTV-T Data Codification and Transmission Specifications for Digital Broadcasting, Volume 4 – GINGA-J: Environment for the execution of procedural applications. São Paulo, Brazil. ISDTV-T Forum (2006)
3. Middleware GINGA Web Site (2008), `http://www.ginga.org.br/`
4. Interactive TV Web Site (2008), http://www.interactivetvweb.org
5. César, P.: A Graphics Software Architecture for High-End Interactive TV Terminals, Helsinki University of Technology (2005) ISBN 951-22-7888-X
6. Bates, P., Atwere, D.: Interactive TV: A learning platform with potential. Learning and Skills Development Agency (2003)
7. John, J.: DVB/MHP Java TV™Data Transports Mechanisms (2002)
8. Moodle Course Management System Home Page (2008), `http://moodle.org/`
9. Walldén, S., Soronen, A.: Edutainment. From Television and Computers to Digital TV (2004)
10. Johnson, K., Hall, T., O'Keeffe, D.: Generation of Quiz Objects (QO) with Quiz Engine Developer (QED) (2005) ISBN: 0-7695-2385-4
11. Heineman, G.T., Councill, W.T.: Book: Component-Based Software Engineering, 1st edn. Addison-Wesley, Reading (2001)
12. MHP Organization Web Site (2008), `http://www.mhp.org`
13. Java TV™ Web Site (2008), http://java.sun.com/products/javatv/index.html

Designing and Developing Process-Oriented Network Courseware: IMS Learning Design Approach

Yue-liang Zhou and Jian Zhao*

Education and Information Technology Institute,
Zhejiang Normal University, Jinhua, Zhejiang, China
Propaganda Department, Wenzhou University, Wenzhou, Zhejiang, China
zhouyl@zjnu.cn, zhaojian@wzu.edu.cn

Abstract. How to design and develop high quality and reusable network courseware has got more and more attention. The existing standard has obtained preliminary achievement. However, IMS Learning Design (LD) can become a new kind of approach. It can transfer the emphasis of network courseware from object-oriented to process-oriented. Based on the introduction of IMS Learning Design, this paper proposed the idea of using LD to design process-oriented network courseware, then a concrete model which use IMS LD tool to develop courseware was provided.

Keywords: IMS Learning Design, Network Courseware, Process-oriented, Learning Activity, Learning Process.

1 Instruction

As the web-based learning and distant education become popular, the network courseware has substituted for traditional stand-alone courseware gradually; it has become the future trend. At present, many computer technology personnel and educators devote to the development of network courseware, but it is still unable to meet teacher's needs. What's the reason? On the one hand, the well-developed courseware is unable to suit for different kinds of instructional strategies; On the other hand, the low level repetition development phenomenon is not avoided.

In order to reduce the waste of resources, at present many research facilities overseas, like IMS, ADL, IEEE, AICC etc, are all devoting to establish the correlative standards. Now some specifications, such as ADL SCORM, IMS Content Packaging, have become mature and obtained preliminary application, but all these specifications pay more attention to content interoperability and re-use, they frequently neglects the instruction process and strategy which are manifested in the courseware.

2 IMS Learning Design and Correlative Technology

From a standards/specifications perspective, IMS Global Learning Consortium has recently released the Learning Design specification (LD,2003), based on "Educational

* Corresponding author.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 52–59, 2008.

Modeling Language" (EML-OUNL,2001), a notational language to describe a "meta-model" of instructional design [1]. As a new international specification, it aims to promote pedagogical expressiveness and reusability of learning content as well, but emphases on re-use and interoperability of learning process and method.

A "learning design" is defined as an application of a pedagogical model for specific objective, target group in specific context. More specifically, it specifies under which conditions, what activities have to be performed by learners and teachers to enable learners to attain the desired learning objectives [2]. The learning design and the included physical resources can be packaged into a 'unit of learning' (UOL). A UOL can be seen as a general name for a course, a workshop, a lesson, etc that can be instantiated and reused for different persons and settings in an online environment.

There are three implementation levels within LD [3], as expressed in table 1: Level A contains the core attributes: people, activities, resources, and their coordination through the method, play, act and role-parts elements. Level B adds greater control and complexity through the use of property and condition. Level C offers the opportunity for more sophisticated learning designs through notification (messaging).

Table 1. The attributes and levels of learning design

LD levels			Attributes	Description
Level C	Level B	Level A	Role	The roles specified are those of learner and staff. Each of these can be specialized into sub-roles.
			Activity	Activities are the connection of role and environment, contain learning activity and support activity
			Environment	Environments contain the resources and references to resources needed to carry out an activity or a set of activities.
			Property	They are used to store information about a person, such as test results or learner preferences;
			Condition	Conditions are If-Then-Else rules that further refine the assignment of activities and environment entities for persons and roles.
			Notification	Notifications are mechanisms to make new activities available for a role, based on certain outcome triggers.

3 The Design Concept of Network Courseware Based on LD

In allusion to the current status, we conceived if there is a method to simple the process of courseware development. According to this method, the courseware can be disassembled into discreteness; the different parts can be managed as object, then

teacher can cooperate with computer professional and other authors, which guarantees the quality of courseware. Like this, the courseware is not merely content stack and technology display. It emphasizes the instruction process and strategy.

The characteristic of LD may realize this conception to design and develop the process-oriented network courseware. First, comparing with other specifications, LD proposed a relatively high-level frame, which guarantees the courseware have general description document and structure. It make the courseware content have a general, simple, easy description and packing method, enables it be applied independently as a logical unit[4]. Secondly, it can provide a run environment to realize courseware frame and content packing. It can develop corresponding application software, achieving the goal of courseware resources sharing and re-use [5]. Finally, it also increased some new characteristics to expand the scope of technical supported instruction method, highlighted the process-oriented characteristic.

The implementation of LD may divide into three steps: creation, production and transmission [6]. In the process, it needs LD tools. They may divide into Editor and Player according to their function. At present more than 20 tools have been developed, such as Reload, Coppercore. Each element of IMS LD can be used to design and develop majority of elements in courseware, its corresponding relations are available in table 2. In the process of developing, LD tool can provide an external concrete environment; moreover the thought behind LD is possible to provide theory guidance.

Table 2. The corresponding elements of network courseware and IMS LD

Elements in network courseware	Property in LD	LD level
personnel	role	A
resource/ material	resource/environment	A
learning strategy/learning flow	Activity structure/activity	A
variable/feedback	property	B
interaction	environment/notification/condition	C
instruction unit	each element in LD	C

4 Development Model of Network Courseware Based on LD

Based on the thought, the author proposed a development model, showed in Fig1.This model is divided into three modules: the courseware development module, management platform, as well as import and output module. These three modules correspond separately with LD implementation three steps, namely creation, production and transmission.

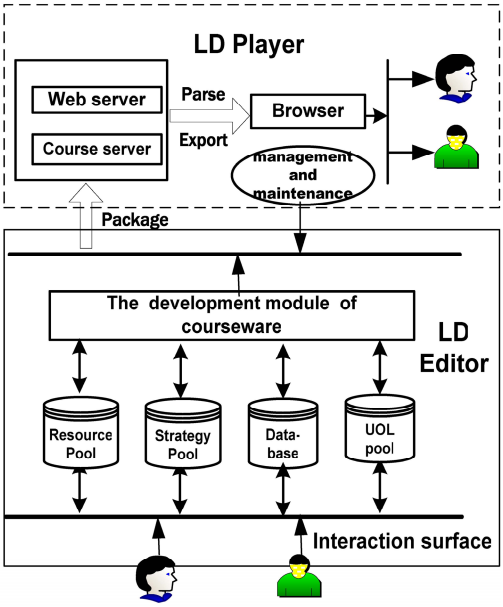


Fig. 1. The development model of network courseware based on IMS LD

This paper chose the Science course in seventh grade (Zhejiang Education Publishing House, China) as an example to expatiate the model. We selected "the science in calendar" of Section 4, Chapter 4, which contains three activities: different calendars, leap year, and the exploration of the twenty-four solar terms.

4.1 The Development Module of Network Courseware

This is the core module in the development process. It provides a dynamic editing environment for teachers to design and develop. The network courseware development module needs to use LD editing tool, like Reload, Collage and so on. . It mainly includes four depositories:

a) Resource Pool

The main function of this pool is providing courseware materials; it corresponds to the learning object (LO) in LD, including many instruction resources and multimedia materials. LD itself doesn't provide tools to develop LO. According to the characteristic of instruction content, need and teachers' own skills, they may chooses many kinds of development tools, like Frontpage, Dreamweaver, PowerPoint. In this example, we used Dreamweaver to develop the activities in the form of Homepage, then put them to LD tool (in this example adopts Reload) to packing.

b) **Strategy Pool**

It is a feature pool of this model that highlights the dynamic developing process. Activity is the center of LD; the learning activity forms the different study flow.

Because the content and learning flow is separate in LD, therefore we also correspondingly separates the resources and teaching strategy in courseware.

A wide variety of pedagogical approaches can be represented by IMS LD, such as problem based learning, game based learning, WebQuest and so on [7]. It will obtain different instructional model and effect by quoting different instructional strategies for the same resources. The utilization of the "activity structure" element may carry on the arrangement、quotation and reference of different activities, which forms their own learning flow. Simultaneously, using the "condition" to choose and decide the next step learning program, thus forms the different learning method and path.

In "the science in calendar", the student's task is mainly to study the three activities, then carry on group cooperation, upload their work; the teacher is responsible to the work: providing feedback, monitoring and managing the entire learning process. According to the three activities, this example adopted the way of sequence. Only completing the preceding activity, can students enter the next step. For example, the teacher provides feedback to students and monitor the learning process, the whole process can be described as Fig2.

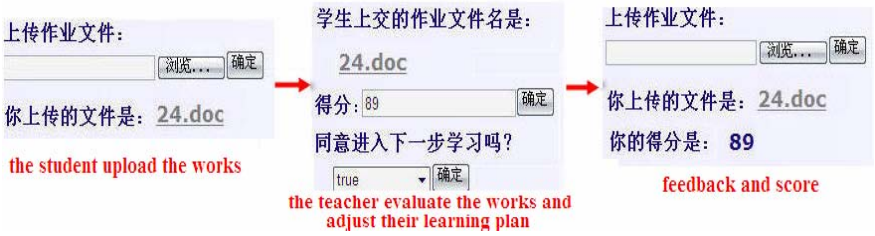


Fig. 2. The teacher's feedback to student's work

c) UOL pool

The courseware material and UOL are both the components of courseware. The material may not become courseware alone, while the UOL can be an independent small courseware, which can be reused.UOL includes the characteristics of courseware, but the courseware is an exceptional case because of some kinds of instruction needs. UOL adopted networking technology, so it can be recombined, which make the teaching resources be shared among different local. The teachers can not only reuse the resources or flow of the courseware, but also the unit module, which enhanced the compatibility and reusability of courseware well.

The specification must enable a learning design to be abstracted in such a way that repeated execution, in different settings and by different persons, is possible. It is necessary to use each element of LD comprehensive to develop UOL that conforms to the request, showed as Table 2. The metadata is used to describe each kind of learning content structure, making it pack as an independent learning object. Each small unit has provided the mutual connection for teachers to retrieval and assembly.

d) **Database**

The database can save other information besides resource and strategy, like the evaluation information, file name, the media type etc. It aims to enable teachers quote the corresponding resources conveniently in the knowledge library. The information in database is temporary, the content renews unceasingly. The attribute of "property" is a variable that can establish the variable and memory in database. It is the essential part of monitoring, personalization, evaluation, user interaction.

In "the science in calendar", we defined three attribute variables: the File variable used for students to upload their work, the Integer variable used to deposit student's learning result, the Boolean variable is defined to decided if enter the next step.

4.2 The Management Platform of Network Courseware

The main objective of this module is to carry on management and maintenance to the overall system. According to the classification of discipline, the nature of curriculum and other specific demands, the developer complete each operation to pools: increasing, deleting and inquiry, providing support for the courseware manufacture. The courseware is deposited in the form of XML documents; simultaneously it needs to examine the consistency, redundancy and validity of database.

LD check and manage the network courseware by using LD player, such as Coppercore,Sled plug.Coppercore is used to check the well-developed courseware, checking its correctness and redundancy, however, the sled plug manage the checked courseware: create, delete and add roles (teacher, student and so on).

4.3 The Export/Import Module

This module aims to realize the share and reusability among different system, providing the function of exporting and importing. It mainly includes two processes: package and parse, simultaneously providing the friendly surface for the user.

1) **Package**

LD platform employs XML language to edit. If all LD systems are expected to share developed courseware, we must pack the courseware. The main usage of LD is to contain learning design inside the content package, thus establishes a learning unit. It did not limit the concrete content and form of resources which packed, but through defining a unified structure to pack different digital resources. Finally, the package is exported in the form of wraps (zip) to server.

2) **Parse and Export**

The result of LD is a XML document. It can be presented as actual learning system by constructing a consistent explanation. After parsed and transacted, the courseware document can be transformed to HTML page document which will be glanced over in the general browser, providing the friendly interaction surface for the learners. In "the science in calendar", the final surface presented to student is showed as Fig 3.

Fig. 3. The interface of "the science in calendar" developed based LD

5 Conclusion

Learning design is one promising new area of both theoretical and technical development. Although LD is not a special specification to develop resource, but it makes the resources development sense, the evaluation and feedback to resource to be more promptly and accurately, enhanced the interaction of resources. At the same time, the work flow of UOL itself is available to share and re-use, which expanded the concept of reusability greatly. In brief, it is a brand-new thought and method to employ LD to design and develop process-oriented network courseware. It causes teacher and the computer technology personnel cooperate well to develop more courseware of high quality, managing and monitoring the learning process well.

However, at present, the resources conform to LD is little. It will take a long time to build a rich resources storehouse. As the LD obtained widespread acceptance gradually, as the technology break through, many current difficulties will be solved as well, more people and organizations will pay attach importance to LD, LD will certainly to receive more welcome.

References

1. IMS Learning Design Specification (2003), `http://www.imsglobal.org`
2. Koper, O.: Representing the Learning Design of UOL, Educational Technology & Society (2004)
3. Sandy: A Review of Learning Design. JISC Project, Britain (2004)

4. Yue-liang, Z., Miao-miao, Z., Gen-qiu, Y.: Developing Educational Paradigm of ICT Application for Supporting Learning Processes and Activities: Learning Design Approach. Mmodern Educational Technology (5), 5–8 (2006)
5. Longkefeng, S.: IMS learning design specification and instruction design (2002), http://www.etc.edu.cn/articledigest30/imsxuexi.htm
6. IMS Learning Design Information Model. Version 1.0 Final Specification (2003), http://www.imsglobal.org/learningdesign/ldv1p0/imsld_infov1p0.html
7. Koper, R., Tattersall, C.: Learning Design: A Handbook on Modeling and Delivering Networked Education and Training. Springer, Heidelberg (2005)

Design and Implementation of Game-Based Learning Environment for Scientific Inquiry

Ruwei Yun, Meng Wang, and Yi Li

Educational Game Research Center of Nan Jing Normal University, China

Abstract. Scientific Inquiry refers to the study that learners conduct with similar or simulated scientific research. It is a problem-oriented learning activity with quite rich contents and forms driven by curiosity. It is also a method by taking the initiative to acquire scientific knowledge and understand science. The game-based learning environment for scientific inquiry (GBLESI) can effectively expand the scope for learners. This article takes the NMR experiments as an example to introduce its Design and Implementation.

Keywords: Scientific Inquiry, VRML, Virtual Environment, NMR.

1 Introduction

The basic characteristics of scientific inquiry study can be summed up as two words—"initiative" and "practice". On one hand, "initiative" refers to learners' ability to act actively, on the other hand, it refers to learners' innovation in the learning activity. Teachers can never forecast all that will happen in the classroom, because learners always display unexpected wisdom which makes them feel amazing. "Practice" refers to learners' real behaviors of observation, thinking, and operation. The objective is to cultivate learners' interest of scientific inquiry, and train their ability of the appreciation of scientific methods, ideas, and spirits, and capability of solving practical problems with scientific methods. In the actual teaching, in order to cultivate learners' ability of scientific inquiry, teachers always make use of the available resources to motivate learners, and provide them with measurable objects, such as the conversion law between ice and water, the usage of solar energy, the relationship between noise and health, and so on. The advantage of available resources lies in its actuality and convenience. However, it has the following three disadvantages. Firstly, it emphasizes more on the activity than learning; secondly, it limits the field of inquiry; and finally, it restrains teachers' ability. The scientific inquiry study under the learning environment of digital games can break the limitation of available resources resulting from problems concerning time, place, capital, and other aspects. It can also eliminate disturbing factors conveniently, and simplify the study content to make it emphasize on the inquiry of the essential character of objects.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 60–69, 2008.

2 Relative Study

2.1 The Bee Dance

The Bee Dance is developed by Illinois University, and the subjects are pupils of the Abraham Lincoln Elementary School. In the fictitious environment, the pupils have no other task except observing and understanding phenomena. In the game, two learners represent two bees who share the same space, and they can put forward questions and think to study the function of bee dance, during which teachers can give appropriate hints.

Fig. 1. Learners in the learn gaming environment for scientific inquiry

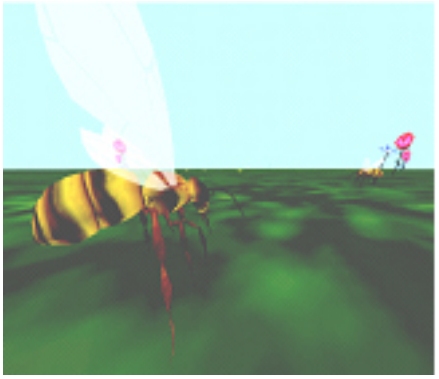


Fig. 2. Virtual environment for demonstrating bee dance

2.2 The Virtual Solar System Project [26]

The Virtual Solar System Project is developed by the education college of Indiana University and the educational technology department of Georgia University. Learners can study the model of solar system themselves in the CAVE environment, and study the movement of the earth and moon through recording trails of them.

Fig. 3. Learners in the CAVE environment to study the models of solar system

3 Design of the Environment of Games for Scientific Inquiry

3.1 Scientific Inquiry Theory

Scientific inquiry, to say simply, refers to the activity that individuals recognize and explain nature and reveal the essence of nature through systemic investigation and study. To understand scientific inquiry exactly, we should know the following several aspects:

Firstly, the objective of scientific inquiry is the nature, and scientific inquiry is to investigate and study the natural phenomena or problems. Through it, individuals can find and reveal the essence and relationships of objects and master the regularity of natural development.

Secondly, as a cognitive activity, scientific inquiry has some procedures and periods. Although science includes many different subjects with different methods and strategies of studies, every study has to go through such similar procedures or periods as formulating problems, hypothesizing, making study scheme, checking hypothesis, making conclusions, and so on.

Thirdly, in scientific inquiry, in order to find and reveal the essence of natural phenomena or objects, individuals have to use a series of scientific methods, such as observation, comparison, classification, measurement, communication, forecast, hypothesis, experiment, making models, and so on. Through these ways, they can find answers to problems about the nature, and know the natural world more deeply. So scientific methods are the soul of scientific inquiry. Of course, different methods can be used in different fields for different problems, and there is no settled or uniform pattern.

Fourthly, as an exploring activity, scientific inquiry has double meanings. It can refer not only to the inquiry in the scientific field, namely, the activities that scientists

do to study the nature, but also to the learning activity of learners, which is similar to the scientific study.

3.2 Scientific Inquiry Process

Firstly, we must stimulate the interests of learners by some circumstances related with issues, which are created by the games. The goal of creating circumstances associated with issues is to expand the thought of the learners, make them put forward their own questions and finally decide what they need to find through the following work. The appropriate game scene for circumstances associated with questions should have the following characteristics:

- Obstacles bringing conflicts, imbalance and the intelligence challenge;
- Interests providing interests, arousing learners to reflect;
- Open solving the problems through wide and various thought, which have no definite answers or solutions.
- Diversity be propitious to learners of all levels to find answers gradually
- Practices searching for solutions of problems by exploration activities. Both individual work and team work is ok.

Secondly, we speculate on solutions of the problems in the game scene. It's a key step for resolving problems and we should remember to consider the learners' age and life experience.

Thirdly, we organize learners to discuss the problems and make plans. This is the preparatory stage for learners to implement activities and also the key points of effective Inquiry. It can be said that the detailed design and feasibility of this step is directly related to the success of the next activities.

Fourthly, we carry out the plan to see if the solution is right. Organizing Learners to discuss and observe the places of the game scenes that we should pay attention to, or introducing them the rules for use and operation, both are important and meaningful for them if they use the equipment for the first time. We also should give learners adequate time to explore, insure that they can finish all the prearranged activities by experimental plans.

Fifthly, here comes the last step to gather everyone together and have their ideas. Learners will communicate with each other and finally educe some common conclusions. Sometimes, work of collecting learners' ideas is carried out in the fourth step also. It is a process of wiping off pseudo things for truth. The results may be positive for our assumptions and may also be negative. Whatever, only when we collect the learners' ideas well and sorted out them effectively, we can make the whole work significative and the final conclusion reflect the truth.

The National Research Council (NRC) of USA published a monograph named Inquiry and The National Education Standards: A Guide for Teaching and Learning. In this book, they expatiate on 5 essential features of classroom inquiry. We can see them in the followed table 1:

Table 1. 5 essential features of classroom inquiry

Features	Expatiation
Question	Learners are engaged by scientifically oriented questions.
Evidence	Learners give priority to evidence, which allow them to develop and evaluate explanations that address scientifically oriented questions.
Explanation	Learners formulate explanations from evidence to address scientifically oriented questions.
Evaluation	Learners evaluate their explanations in light of alternative explanations, particularly those reflecting scientific understanding.
Communication	Learners communicate and justify their proposed explanations.

4 Design of Game Based Learning Environment for Scientific Inquiry on NMR Experiment

4.1 Framework and Process Design

The virtual environment is designed for single person with 3 main parts（As is showed in Fig4 below）: Show pictures of NMR CT, Scan human body with the NMR CT System, and Reveal essence of The NMR.

The first part is used to show pictures made by NMR equipment (e.g. NMR CT pictures of the head, NMR CT pictures of the backbone, NMR CT pictures of the chest and so on.). In this part, we can give out both the healthy organ pictures and bad ones at the same time, so we can see their differences. This part is used to arouse the interests of learners and make it a proper time to introduce the NMR theory. Teachers will give some questions to learners in this part.

The second part is used to simulate NMR CT System. In the virtual environment, we create several persons and apparatus to simulate the real activities in the hospital. Learners interact with the models and get proper response: movies, pictures, words for explanation and so on. By doing research in the environment, learners will gradually summarize the condition for NMR phenomenon. For the whole job of creating such a virtual learning environment, the second part plays the most important role. It must offer learners enough chance to interact with the environment and get plenty of proper response. In order to keep the learners' interests, we have to make the environment somewhat like a game, such as setting up an area to display the learners "Scanning points" which is decided by their activity in some scenes: configure the apparatus, change the water content in human body, scan more places and answer more questions .

The last part will make learners close to the essence of NMR theory in forms of movies, pictures and text. After gathering the learners to discuss their research and the phenomenon in the second part, teachers can use this part to make them understand the theory better.

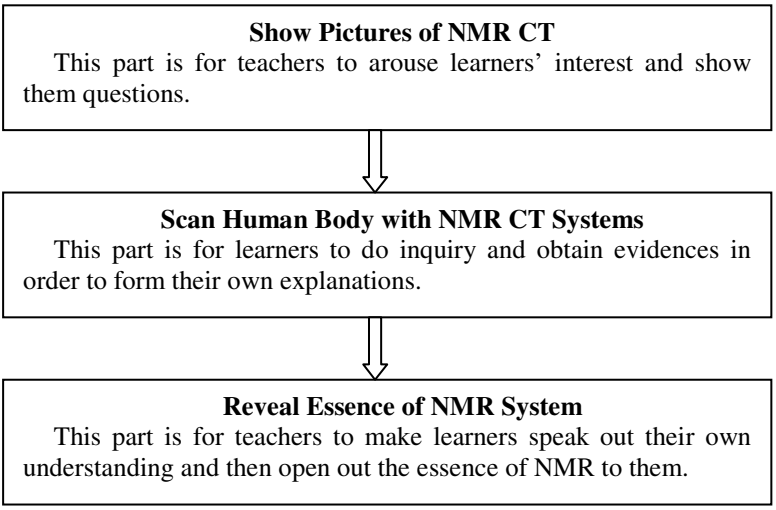


Fig. 4. Modle of the Frame work and process

Process design is as follows:

(1) Teachers use the first part of the environment to introduce the conception of NMR, they show some pictures and then ask questions about them. The teachers also have to introduce the environment to learners, help them enter it and guide their research journey.

(2) Learners open the second part with questions discussed before. They manipulate the models in the way their teacher have told them, including choose characters, change viewpoint, press switch buttons and so on. By interacting with the environment and getting responses, they gradually form their own idea about the questions. Detailed steps are as follows:

Step1: Press "Scan human body with the NMR CT System" button to enter the second part. There are two characters in the scene that named audience A and audience B. We call them A and B as a short name. Choose A or B. If A is chose, he will lie down on the board and let B as the doctor to control the panel. So the same if audience B is chose first.

Step2: Suppose A is chose first, he press "Ready" button and then be sent into the inside room of NMR CT apparatus if audience B press the "Close Board" button on control panel.

Step3: Two buttons will come out to let learners choose the viewpoint. If they choose A, the scene will not change and B will control the panel to send A into the scanning device. If B is chose, scene will change into the inside of the device. Now the control panel will under A's control.

Step4: Whatever the viewpoint is, learners will get the same group of buttons, they are: Alternating Magnetic Field ON/OFF, Radio Frequency Signal On/OFF, Increase/Reduce water Content in Human Body (15%), Scan Head, Scan Chest, Scan Backbone, Scan Legs. Learners press some of the buttons to set up the device and

scan the human body. Now the Scanning Simulating Scene will be loaded into the environment.

Step5: Inside the scanning device, we use some digital and laser light effect to simulate the scanning process of the magnetic field and the radio frequency signals.

Step6: Both inside and outside of the scanning device we set displayers to display the result of learners' work, that usually in forms of pictures, movies and score.

Step7: Learners press buttons to change the magnetic field and the radio signals, they scan many places of human body and get different feedback. By doing research work in the virtual environment, learners will gradually find there own idea about NMR CT System's working condition.

Step8: At the top left corner of the scene, we set an area to show the learners' "scanning score", which is decided by how they set up the device, how many places they scanned and how they answered the questions. This part is the embodiment of the environment's challenge and fun as a "game based " learning environment.

(3) After most learners have finished their research work, teachers gather them together to talk about the questions. Teachers should listen to every speaker's words carefully and make proper comments. At the same time, teachers should better use the environment again when they make comments or explain some theory.

(4) At last, teachers use the third part to explain the formal and in-depth theory of NMR. It is a nice choice to show the learners some movies here.

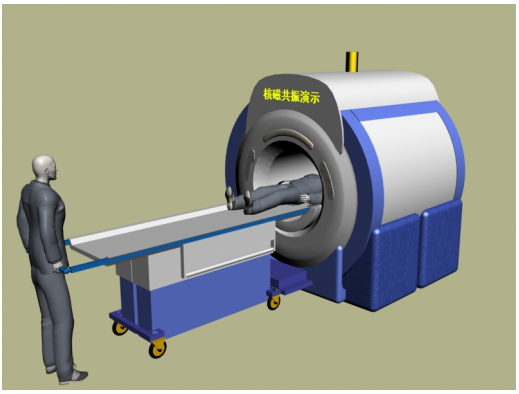


Fig. 5. Scene of transporting audience A into NMR CT apparatus

4.2 Realization Techniques

Creation of the Virtual Environment

We choose 3DS Max8.0 to create complicated models and animations in the virtual environment, for instance, outward appearance and inner structure of NMR scanner, animation to simulate the scanning process or to simulate the rotation of nucleus in magnetic field. As the professional software of creating/rendering 3D model and animation, 3DS Max can make very fine objects to attract learners' attention. We also use a powerful plugin of 3Ds Max named Polygon Cruncher 7.22 to optimize the

models. This small tool will not destroy the model's material quality and details while simplifying its surface. After being optimized, models become smaller and easier for explorer to load and display. The more simple models are, the more smoothly for explorer to change the scene. We choose VRML2.0 to create the virtual environment world and organize the models. Some simple models (such as wireframe, box and sphere) and animations (such as translation and rotation) are also created by VRML using its inner node.

We use the inner node of VRML2.0 to create the most static scene and its transform. In order to improve code efficiency and reusability, we need to create many new nodes, which have their own fields and functions. Here the object-oriented thinking is necessary , the new nodes should be encapsulated and inheritable. So we can create nodes based on the proto mechanism.

Proto mechanism includes three parts: declaration, implementation and instance. The keyword proto is used to declare the node type defined by developers. The definition of a node type contains field, exposedField, eventIn, eventOut and nodebody. Proto can be defined as:

```
PROTO<node type name>[<interface list >)]{<node body>}
```

LOD is a system feature for optimizing the amount of details rendered in a scene. VRML provides an LOD node, which can explicitly change different leveled detailed versions of model. We employ this way to optimize VRML files so as to accelerate rendering speed in real time walkthrough.

Control of the Virtual Environment

The VRML language can only do simple interaction in the scene. In order to improve this deficiency, company SONY has raised a solution called JSAI (Java Script Authoring Interface) [11], which has realized advanced script function of VRML and provide a new way to control objects in VRML. JASI consists of three packages: vrml、vrml.field and vrml.node.

vrml package, which contains field, ConstField, MField, ConstMField, Browser, Event and BaseNode;

vrml.node package, which contains script class and node class. These two classes are inherited from vrml.BaseNode class;

vrml.field package contains various classes to define data type of fields in VRML. These data types are used to describe position, rotation, and time and so on. To every data type, JSAI provides a special class to describe. Browser navigates according to the url field of script node. For example,

```
        Script{
              url"../scripts/javaScript.class"
  }
```

When a browser loads java script (e.g. javaScript.class) successfully and the script node receives a group of eventIns, vrmlscript.class will call the method ProcessEvents() of script node automatically. To every event object, getName(), getTimeStamp() and getValue() can be used to get the name of the eventIn, the time when the event is sent and the value of the event respectively.

Script program can read and write the Fields defined by users in script nodes through their field names, and the method getField() provide a way to refer interface fields.

4.3 Challenges and Other Techniques

Challenges we still faced

One problem we still faced is how to balance the game content and knowledge content in creating such a game-based learning environment. How to make the game content at a proper level so that learners can not only enjoy the inquiring journey, but also keep enough attention on the questions they met before? This problem is really concernful. Perhaps the solution lies between the research of learners' psychological characteristics and making the game content alterable for users at different ages.

The other challenge is the inconsistency between its complexity and acceptability for schools in china. Generally speaking, the more complex the simulated environment is, the better effect it can provide. Considering the experience quality of learners getting from the inquiring journey, we should make the virtual environment more complex and absorbing using some powerful game development kit (such as Microsoft XNA Game studio) or some game engine, other than VRML. However, we need to be wake up to the realities in china. The most proper place to use such virtual environment is classrooms in middle school. If we want our environment fully play its role, we must make sure that both the teachers and computers cater to the operation of the environment. Unfortunately, computers in most schools of china are not good enough to ran complex 3d games well, which need powerful display card to render its fascinating effects. Programs ran in internet explorer with a little plugin maybe easier for middle teachers to use than complex 3d game-base learning environment.

Other techniques we may use in the future

The Java 3D API is an application programming interface created by Sun Microsystem. It's a very powerful language for writing three-dimensional graphics applications and applets. It gives developers high-level constructs for creating and manipulating 3D geometry and for constructing the structures used in rendering that geometry. Application developers can describe very large virtual worlds using these constructs, which provide Java 3D with enough information to render these worlds efficiently. Compared with VRML, Java 3D API is more powerful in creating large virtual worlds with lots of objects and interactions. However, if the virtual world we planed is not very large and interactions we designed are not quite complex, the VRML will be a better choice than Java 3D.

OpenGL is a library of functions for 2D/3D virtual world rendering. Besides its basic library, there are also some additional libraries to make the program work easier, such as Glu, Glaux, Glut and Glx libraries. It is an industrial standard proposed by SGI and some other companies, so it can perform well on any kind of computer hardware and system. Unlike DirectX of Microsoft, OpenGL is free for everyone to use and develop his own product. The programming language C++ can work well with OpenGL and together they can make more exciting 3D application with high efficiency than Java 3D. However, Programming in C++/OpenGL is not as easy as VRML and Java 3D.

5 Conclusion

Compared with the scientific exploration in real life, game based learning environment for scientific inquiry avoid the constraints of space-time, security and resources. It can lower the cost to provide learners relatively rich learning content. Incorporate appropriate games, problem –oriented virtual environment can stimulate the interest of learners well and provide them with joy and rich learning activities.

There are varieties of ways to implement a game based learning environment for scientific inquiry, In this article, we introduce a new way to implement the virtual environment for NMR experiment with VRML2.0, JavaScript and 3DS Max modeling tools. Through practice test, we have achieved good results and hope it useful for others' design or implementation.

References

1. Downes, S.M.: Truth, Selection and scientific inquiry. Biology and Philosophy 15(3), 425–442 (2000)
2. Davis, K.S., Falba, C.J.: Integrating Technology in Elementary Preservice Preservice Teacher Education. Orchestrating scientific inquiry in Meaningful Ways 13(4), 303–329 (2002)
3. Ieronutti, L., Chittaro, L.: Employing virtual humans for education and training in X3D/VRML worlds. Computers & Education 49(1), 93–109 (2007)
4. National Research Council of USA: Inquiry and the National Science Education Standards: A Guide for Teaching and Learning, National Academies Press (2000)
5. Murano, P., Mackey, D.: Usefulness of VRML building models in a direction finding context. Interacting with Computers 19(3), 305–313 (2007)
6. Liu, Q., Sourin, A.: Function-based shape modelling extension of the Virtual Reality Modelling Language. Computers & Graphics 39(4), 629–645 (2006)
7. Koročsec, D., Holobar, A., Divjak, M., Zazula, D.: Building interactive virtual environments for simulated training in medicine using VRML and Java/JavaScript. Computer Methods and Programs in Biomedicine 80, 61–70 (2005)
8. Moore, K., Dykes, J., Wood, J.: Using Java to interact with geo-referenced VRML within a virtual field course. Computers & Geosciences 25(10), 1125–1136 (1999)
9. Whiteman, P.: How the bananas got their pyjamas: A study of the metamorphosis of preschoolers' spontaneous singing as viewed through Vygotsky's Zone of Proximal Development. A Thesis for Doctor Degree of Philosophy, the University of New South Wales (2001)
10. Yun, R., Pan, Z., et al.: An Educational Virtual Environment for Studying Physics Concept in High Schools. In: Lau, R.W.H., Li, Q., Cheung, R., Liu, W. (eds.) ICWL 2005. LNCS, vol. 3583, pp. 326–331. Springer, Heidelberg (2005)
11. Davison, A.: Pro Java™ 6 3D Game Development, pp.3–12. Apress (2007)
12. Dave, A., Kevin, D.H.: Beginning OpenGL Game Programming, Course PTR (2004)
13. Multigen-Paradigm Incorporated (online2006.01.20), `http://www.multigen.com`
14. James, G., Bill, J., et al.: The Java Language Specification (1996)(online2006.01), http://java.sun.com/docs/books/jls/html/index.html

Research and Implementation of Web-Based E-Learning Course Auto-generating Platform

Zhijun Wang, Xue Wang, and Xu Wang

Educational Technology From Tianjin Normal University Tianjin, China
army.w@163.com

Abstract. Aiming at the actuality that it is difficult and high-cost to teachers whose professions are non-computer to develop web-based E-Learning courses in traditional way, the research developed a platform of auto-generating web-based E-Learning courses by adopting systemic and scientific method and technology of ASP.NET 2.0. In the platform, through some simple operations, for instance, selecting function modules of teaching activities and editing teaching contents according to instructional strategies, an independent web-based E-Learning course which is in accordance with characters of students and subject and teaching regulations of E-Learning can be generated and the instructional design philosophy of teachers can be realized. The web-based course generated by the platform can be maintained easily and updated again and again.

Keywords: E-Learning, auto-generating, web-based course, instructional strategies.

1 Introduction

Web-based course as the main approach to carry out teaching activities of E-Learning plays an important role in E-Learning, thus courses of high quality are the important guarantee to develop E-Learning. However, the current status of developing web-based E-Learning courses is not optimistic and the courses that can be used in practice are not the majority. There are several problems: first, most of persons who develop web-based E-Learning courses are members of computer profession, but teachers who truly understand characters of students and subject and teaching regulations of E-Learning have serious difficulty in developing web-based courses. So the effect of teaching activities is discounted and teaching design can't be realized. Second, although the forms of presenting teaching contents and carrying out teaching activities are various, it is difficult to update teaching contents and activities according to characters of students and subject and teaching regulations of E-Learning [1]; Third, because of the differences of running environments and the independence of development, the modes of running and maintaining are not united [2]. The problems hereinbefore influence the quality of web-based E-Learning courses severely. On a long view, low-quality web-based E-Learning courses not only waste time, money and human resource of developers, but also lead the result that students gain little and teachers loose

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 70–76, 2008.

interests [3]. Therefore, the advantages of E-Learning are weakened and the development of E-Learning is baffled. In order to settle the problems above, the research developed a scientific and valuable platform of auto-generating web-based E-Learning courses. After being trained, teachers who truly understand characters of students and subject and teaching regulations of E-Learning can develop web-based E-Learning courses without the help of members of computer profession. The courses include instructional design philosophy of teachers and can be maintained and readjusted anytime and easily.

2 Design of Platform

2.1 Basic Ideas and Principles

A web-based E-Learning course is the confluence of teaching contents and activities of some subject [4]. The teaching contents and activities are presented by network. So a web-based E-Learning course can be divided into two parts: teaching contents and web learning and teaching environment organized by teaching objectives and strategies [5]. Teaching contents of web-based E-Learning course can be presented by web page, teaching activities can be realized by BBS, online test and so on. Aiming at the two parts of web-based E-Learning course, the platform provides function modules of teaching activities and teaching contents (have interface of editing contents). Through some simple operations of teacher, the diverse function modules of teaching activities and teaching contents have been edited can be composed to a web-based E-Learning course of systematization and independence. As the platform adopted B/S (browser/server) mode, the clients need not to install any software, after logging in the platform through browser, teachers can develop courses online. The platform emphasizes easy and humanistic operation and functions available for selection among diverse modules and within one module. Because the platform provides instructional strategies templates, under the guidance of the platform, teachers who aren't familiar with teaching philosophy can design and develop web-based E-Learning courses in accordance with characters of students and subject and teaching regulations of E-Learning.

2.2 Function Design of Platform

The platform designed 14 function modules. Function module of course development was primary and others were secondary. The function design of the platform was based on two identities (Fig. 1).

Identity of Administrator

- User Management: Manage users who develop courses through the platform. The administrator can query, delete and create users.
- Course Management: The object is all courses developed by the platform. The administrator can query, delete, edit and create courses.
- Resource Management: Manage the resources uploaded by users. Create and delete resource categories. Query, delete and upload resources.

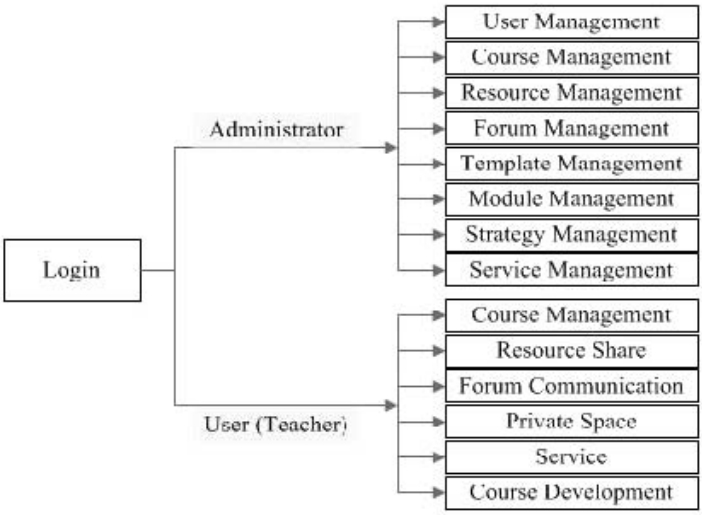


Fig. 1. Function design of platform

- Forum Management: The object is module of Forum Communication. Create, delete and edit boards, topics, bulletins, messages and so on.
- Template Management: Manage templates of courses provided by the platform. The administrator can create, delete and edit templates of courses.
- Module Management: The object is function modules available for web-based E-Learning course development. The administrator can query, create, delete and edit the modules.
- Strategy Management: Manage the instructional strategies provided by the platform. The administrator can query, create, delete and adjust instructional strategies.
- Service Management: Query, add, delete services provided for users.

Identity of User (Teacher)

- Course Management: The object is the courses developed by users. User can query, delete and edit courses developed by him/her.
- Resource Share: Users can upload and download all kinds of resources that can be used in their courses.
- Forum Communication: The platform provides a space for communication. Users can communicate the experiences of course development with others through it.
- Private Space: View and update private data and password, write working log and so on.
- Service: Provide FAQ, guidance and help for users.
- Course Development: Users can create, preview, edit and publish courses.

2.3 Design of Course Development Module

The module is the core of the platform. The flow (Fig. 2) of course development is: after applying a new course, user chooses template of course's style and instructional strategy to confirm function modules of course (user can design function modules if he/she don't choose instructional strategy) and enter the interface of course making. In the interface, user can edit function modules, page contents and style of the course. The course could run as an independent web-based E-Learning course after published.

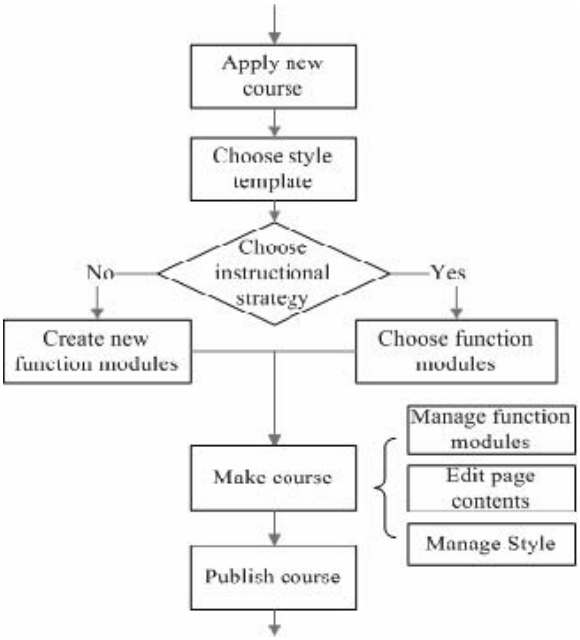


Fig. 2. Flow of course development

In order to adapt the teaching regulation of E-Learning, the platform emphasizes building learning and teaching environment as well as the course contents. 19 optional function modules of teaching contents and teaching activities divided into two groups are provided, they are:

The function modules group of teaching contents includes 9 modules: teacher introduction, course introduction, teaching objectives, classroom video, teaching outline, teaching scheme, electronic teaching material, network courseware and teaching plan.

The function modules group of teaching activities 10 modules: experiment instruction, homework & exercise, online test, evaluation & feedback, teaching blog, teaching interactivity, students'space, outcome display, relevant resource, help & guidance.

User can choose function modules of course according to need or design new function modules.

The platform provides instructional strategy templates to help user to develop courses in accordance with characters of students and E-Learning. The instructional strategy templates of the platform followed the instructional design principles as follows: Pay attention to the analysis of the teaching objectives and teaching contents; Pay attention to create circumstances; Pay attention to the design of all kinds of information resources; Emphasize teachers' guiding role and students' participating role [6]. The platform provided 4 instructional strategy templates (each instructional strategy template includes a series of initial function modules that are in accordance with characters of the instructional strategy), they are: **Information delivery mode.** The characters of the instructional strategy are publishing teaching information by web to show teaching plan, teaching contents and providing relevant resources. The mode includes 6 function modules: teaching objectives, teaching outline, teaching plan, electronic teaching materials, homework & exercise and relevant resources. **Discussion mode.** The characters of the instructional strategy are assisting learner to learn and providing services such as guidance and discussion [7]. Learners attain learning goals through group discussion and study. The mode includes 7 function modules: teaching objectives, electronic teaching material, teaching blog, teaching interaction, relevant resources, help & guidance and evaluation & feedback. **Collaboration mode.** The instructional strategy stresses intercommunion activities between students. The mode carries out co-operative learning by creating supportive environment. Make students form a learning community by carrying out co-operative activities such as question, answer, resources share between them. The mode includes 7 function modules: teaching objectives, electronic teaching material, teaching interaction, relevant resources, help & guidance, outcome display and evaluation & feedback. **Information synthesis and creation of resources mode.** The instructional strategy is to realize the assimilation, interaction and synthesis of information resources by student's active finding, creating, organizing and reorganizing the contents of concrete knowledge domain [8]. The mode includes 6 function modules: teaching objectives, teaching blog, relevant resource, students'space, help & guidance and outcome display.

User can choose the appropriate instructional strategy, create new instructional strategy or alter the original instructional strategies to answer for characters of students and E-Learning. The instructional strategy templates of the platform are of great value and reference for instructional design of web-based E-Learning course. Therefore, the quality of web-based E-Learning course can be guarantied.

3 Implementation of the Platform

3.1 Development Tools

The platform used ASP.NET2.0 as its key technology. A series of new features of ASP.NET2.0 such as Master Page, Website Navigation, User Management, Profile and Theme/Skin were applied to enhance the efficiency and quality of the platform development. The software development environment of the platform was Visual Studio 2005 and its background database was SQL Server 2005.

3.2 Architecture of the Platform

According to the idea of software engineering, the architecture of the platform was composed of four layers so as to achieve code reusablity, maintainability and expansibility [9]. Fig.3 shows the four layers architecture of the platform.

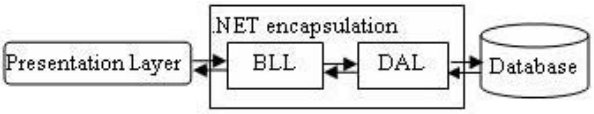


Fig. 3. Four layers of Platform

Presentation Layer: Presentation layer was composed of ASP.NET pages that share a common page layout.

Business Logic Layer (BLL): Business Logic Layer enforced custom business rules.

Data Access Layer (DAL): Data Access Layer served as bridge between BLL and Database Layer.

Database Layer: Database Layer was used for storing data.

In four layers architecture each layer only interacts with its neighbor layers. Using well defined interface, the inner implementation of each layer was irrelevant to others and switch of heterogeneous database can also be achieved easily by updating configuration files. Thus, the efficiency and quality of the platform development can be enhanced.

4 Conclusion

The purpose of the research is to develop a platform of auto-generating web-based E-Learning course, in order to make convenient for teachers whose professions are irrelevant to computer to develop web-based courses that follow characters of students and subject and teaching regulation of E-Learning. The platform applied teaching philosophy to function modules and used guide pattern to instruct teachers to develop courses, so the quality of courses can be guaranteed. The platform is using in our university now. It is approbated by teachers and students. Teachers and students'enthusiasm for E-Learning is greatly enhanced and the teaching quality of web-based E-Learning courses is improved. The platform has promoted the further development of E-Learning to some extent.

References

1. Luo, H.: Actuality and Countermeasure of Network Courses Development. Modern Distance Education 4, 33–34 (2003)
2. He, K.: Modern Educational Technology and Design and Development of High-quality Network Courses. China Audiovisual Education 6, 5–310 (2004)

3. Chen, Y.: Research and Improvement of Course Development Tools Based on Web, pp. 3–8. Beijing Normal University (2002)
4. Schewe, K.-D.: A Conceptual View of Web-Based E-Learning Systems. Education and Information Technologies, 81–108 (2005)
5. Niemi, H., Nevgi, A., Virtanen, P.: Towards self-regulation in Web-based learning. The annual meeting of the American Educational Research Association (2003)
6. He, K., Li, K., Xie, Y.: Theory Fundament of Guide-Participant Teaching Mode. E-Education Research 2, 3–9 (2000)
7. Cui, M., Zhang, R.: Application & Research for BBS in Online Education. Journal of Chang Chun Teachers College 23(3), 43–45 (2004)
8. Beck, I.L., McKeown, M.G.: Inviting Students into the Pursuit of Meaning. Educational Psychology Review 13(3), 225–239 (2001)
9. Cazzulino, D.: Beginning C# Web Applications with Visual Studio.NET, pp. 223–470. Tsinghua University Press (2003)

A Humanized Mandarin e-Learning System Based on Pervasive Computing

Yue Ming, Zhenjiang Miao, Chen Wang, and XiunaYang

Institute of Information Science,
Beijing JiaoTong University, Beijing 100044, P.R. China
myname35875235@126.com

Abstract: This paper presents the design and implementation of an application system where pervasive computing is used in Mandarin e-Learning. First a short introduction to the ongoing changes concerning pervasive computing and particularly e-Learning is given, followed by an e-learning framework including the application server and enterprise server, which distributes the personalized learning services to the learners so that the learners can obtain the learning resource or personalized learning guide interactively. Its design and implementation issues are discussed in details.

Keywords: Pervasive computing, Mandarin e-learning, MVC, database.

1 Introduction

With Chinese opening and fast economy development in recent years, the communication between China and the world becomes more and more important in a wide range. So Mandarin, the important communication tool and culture carrier that lets the foreign countries know China, attracts more and more governments, educational organizations and corporations. Consequently, hundreds of Mandarin learning platforms, tools, electronic learning materials, and so on, have been developed, such as Go to China Mandarin Platform Advance in 2004 [1], StepByStep by HeiLongjiang University in 2001 [2], Chinese Master by Yuxing Soft LTD. in 2004 [3], etc. Although they are all good tools for learning Chinese and have been designed to help students to learn a large number of Chinese words as quickly as possible, they only shift the problem to the dependency on an available and appropriate disc drive or software.

The spectacular development of Internet provides an untraditional and wide way of education. As a result, e-Learning is becoming a fixed part of most people's life, as they are forced to life-long learning, no less than our Mandarin Studying. The e-Learning system can also afford communications between the foreign learners and the teaching server. There are some typical Mandarin e-Learning systems, for example Chinese Horizon Mandarin Training by Yahoo in 2004 [4], The Ottawa Mandarin School's curriculum by Taipei language Institute in 2001 [5], EASE Mandarin by Mandarin House Language Institute in 2001 [6], etc. However, our investigation indicates that the teaching systems of e-Learning are lack of intelligence. They cannot

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 77–87, 2008.

interact with learners, cannot adjust the curriculum contents based on their learning situation, and most importantly cannot make the learning anywhere and anytime.

Pervasive Computing emerges as the times requires and gradually penetrates our daily life. When we talk about the main advantages of pervasive computing, we usually think of anywhere, anytime, any format, and any device [7]. This means:

- Anywhere: global accessibility, with regard to various kinds of communication networks.
- Anytime: twenty-four hours, but also independent of other services or persons .
- Any format: email, public services, inter- and intranet, various data formats.
- Any device: (Table-) PC, Personal Digital Assistant (PDA), cell phone, etc.

Up to now, the computing scientists have developed a wide range of applications in pervasive computing, for instance, pervasive retailing information systems [8], wide-area e-health monitoring [9], domestic ubiquitous computing applications [10], etc. But pervasive computing based language e-Learning systems, especially Mandarin e-Learning, have not seen so far. Thus we can take full advantage of pervasive computing technologies to design our Mandarin e-learning system where the learner can put forward his personalized learning requests according to his knowledge structure and learning plan. Then, our system will analyze the learner's learning history and demands, adjusting the curriculum contents, which stand for his learning requests and preferences anytime and anywhere without any restrictions.

Moreover, in order to facilitate the foreign learners and bring more and more web-based Mandarin e-learning resources and personalized learning guides, we design pronunciation component as a major function of our system where Mandarin language learners can practice their pronunciations including tones. In a word, our system is a pool of Mandarin language place where you can interact with people from all around the world to share your interests and concerns about Mandarin.

The structure of this paper is as follows: our e-learning system is described briefly in section 2; in section 3 and section 4, the application server and the enterprise server are presented separately: then we describe the core function of our system—pronunciation assessment in part 5; finally, the design and implementation of our Mandarin e-learning system as a whole is introduced in details.

2 System Description

In terms of the concept of pervasive computing discussed above, we design our pervasive Computing Mandarin e-Learning System as in Fig.1.

To implement this system, Java is used due to the isomerism of platform, and JSP used to design the web page. We put the web server and IP-to-PSTN gateway in two computers. Apache Tomcat is used for the web server and Asterisk used to provide a central switching in the gateway [11].

The Web Server in Fig.1 is further illustrated in Fig.2, as it is the core of our humanizing e-Learning system. It is composed of two main parts—application server and enterprise server. The Teacher Agent in Fig.1 provides a major e-Learning function of the system — pronunciation evaluation. They will be described in following sections in details.

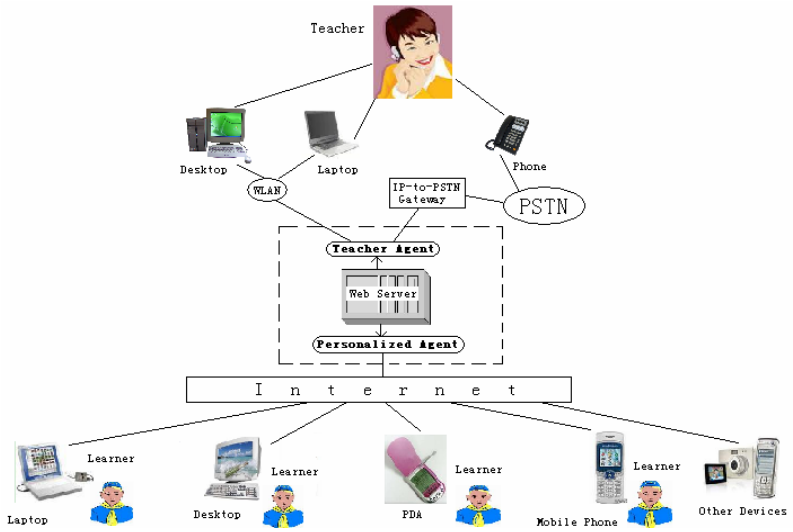


Fig. 1. Pervasive Computing Mandarin e-Learning System

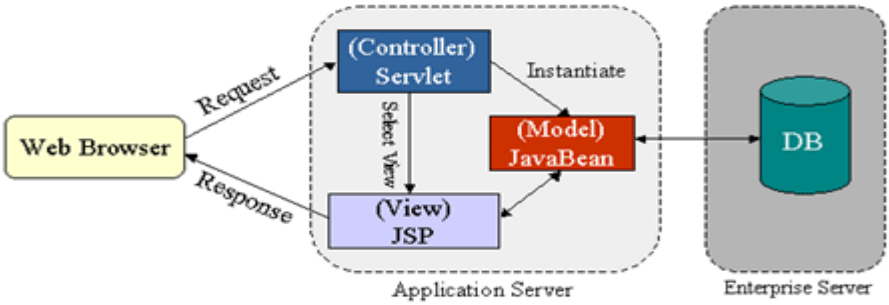


Fig. 2. The Framework of our Web Server

The prominent advantage of our system is to teach students in accordance with their aptitude (humanizing) and condition with flexible time and space. The system also facilitates communication between the teacher and the students (e.g., guidance during students' work on assignments) [11].

3 Application Server—MVC (Model-View-Control)

Our System is a user-friendly, web-based information system for the analysis of students from population studies. Users can access their courses from all over the world simply through a web browser at any time.

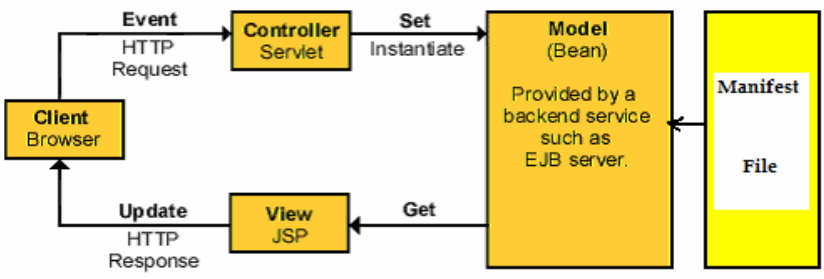


Fig. 3. The Framework of Application Server

We consider an idea of how the learning material is organized in the Fig.3. It consists of two major components: A so-called 'Manifest File' and the resources (the physical files with the actual learning contents). The Manifest File describes the (hierarchical) course organization in XML. Each manifest consists of a metadata section (information about the course, student's preference, and mental states, etc.), an organizations section (the structure of the course) and a resource section references to the physical files) [7].

1. The Model on the server side processes the original manifest and the contents. It produces a manifest with additional information (size of the whole course, size of the contents, the audio data, etc.).
2. The Model knows which courses are available and where these are located. When Controller receives a HTTP request from Client and instantiates it to the Model, the Model sends the manifests and corresponding contents in question to the View.
3. Finally, the View is responsible for updating the HTTP response accordingly (the course structure and the actual contents of the course material) to Client.

The first step in the process is converting the manifest and the contents to a viewer conformant manifest. Therefore, the Model was developed, which modifies the original manifest. The reasons for these modifications are to provide the View with all required information. Consequently, this leads to an accelerated data calculation and transports to the Client and also allows the user to download whole courses from the Controller. Thus, the user can concentrate on specific contents and store just the corresponding course material on the Client, leaving out the rest. Once documents are not of interest anymore, they can be deleted without having to delete the whole course.

The Model represents the server side of the process, connecting to the Client application. The core tasks of the Model are:

- Registration of courses (distributing the available course based on the learners studying level)
- Communication with Client (file transfer via TCP/IP)
- Logging (of the HTTP request events)

By adding a manifest to the Model a course becomes available for download to View. The View can connect the Model and HTTP response. The View sends a list of all available courses to the Client (all courses that were transformed into an appropriate manifest and registered with it).

The last element in the presentation process is the View. It is the application the learner is using, whereas the course developer or providers use the Controller and the Model, respectively. The View features are: showing available course, loading and deleting course manifests, the presentation of the course structure, etc. If the user is no longer interested in some chapter anymore, they also can delete the responding contents by selecting the individual contents or whole chapters.

4 Enterprise Server–Database

Database management information and analysis system is a novel system that integrates a friendly web-used user interface a number of information required by students and teachers [12]. It is composed of three major subsystems: user logic module, e-Learning module and user information module.

User logic module is the component that manages the learners' information stored in the database [12]. After registering, the students can login our e-Learning system by entering the users' name and ID as in Fig.4. If some student enters our system, a personalized agent will be invoked. The agent knows learner's studying history. It can help to find a chapter that satisfies what he needs automatically and improve the learning quality greatly [11].

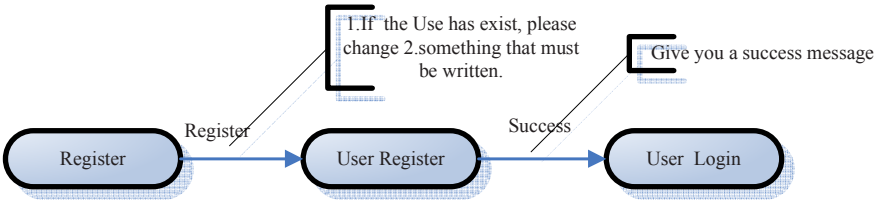


Fig. 4. The Flow Chart of User Logic Module

E-learning module has been designed so that it can be customized to a user's particular demands without a great effort, especially the humanized lessons' contents.

• For beginner, our system will deliver prepared easy content to students, and ensure that the content reaches them anytime no matter where they are.

• For outstanding learners with great fluency, the evaluation was executed with a brief quiz offered at the end of the lecture.

• If the learner feels tired, the system can make studying more interesting by numerous means like hearing/seeking someone, arranging a cartoon or game, etc.

Table 1. The Table of students

Table	T_ Students			
Column	Description	Type defination	null/not null	Restriction
UserId	User id	int IDENTITY(1,1)	Not null	key
UserName	User name	Varchar(20)	Not null	
PassWord	User password	varchar(50)	Not null	
realname	User real name	varchar(20)		
Sex	User sex	varchar(2)		
Address	User address	varchar(150)		
nationality	User nationality	varchar(50)		
post	User post	varchar(8)		
telephone	User telephone	varchar(25)		
Email	User Email	varchar(50)		
RegTime	User regtime	datetime		DEFAULT(getdate())

键	ID	名称	数据类型	大小...	空	默认值
	✓	userid	int	4	☐	
		username	varchar	20	☐	
		password	varchar	50	☐	
		realname	varchar	20	✓	
		sex	varchar	10	✓	
		nationality	varchar	50	☐	
		address	varchar	150	✓	
		telephone	varchar	25	✓	
		post	varchar	8	✓	
		email	varchar	50	✓	
		regtime	datetime	8	☐	(getdat

Table 2. The Table of Lessons

Table	T_Lesson			
Column	Description	Type defination	null/not null	Restric-tion
Lesson_id	Lesson id	int IDENTITY(1,1)	Not null	key
Lesson_name	Lesson name	Varchar(50)		

键	ID	名称	数据类型	大小...	空	默认值
	✓	lesson_id	int	4	☐	
		lesson_name	varchar	50	✓	

User information module is used to preserve the Personality Factors. It contains dimensions such as Personal, Preference, and Portfolio, etc. They are coupled with extension features so that distributed e-learning systems are the summation of mentality requirements, which affected not only the style of interaction, but also the style of behavior in learning (speech and tone). This module also deals with editing/deleting users or lessons depicted in Fig.5.

Moreover, our system also has a security schema that supports two types of user roles. The administrator role, which grants full access to all features of the program, and the simple user role, which grants limited functionality. This limited functionality

Table 3. The Table of Contents

Table	T_Content			
Column	Description	Type defination	null/not null	Restric-tion
content_id	Content id	int IDENTITY(1,1)	not null	key
lesson_id	Lesson ID	int not null	not null	
chinese	Chinese content	varchar(100)	not null	
phoneticize	Chinese phoneticize	varchar(100)		
english	English	varchar(100)		
pronounce	Pronunciation	varchar(100)		

键	ID	名称	数据类型	大小...	空	默认值
	✓	content_id	int	4	☐	
		lesson_id	int	4	☐	
		chinese	varchar	100	✓	
		phoneticize	varchar	100	✓	
		english	varchar	100	✓	
		pronounce	varchar	100	✓	

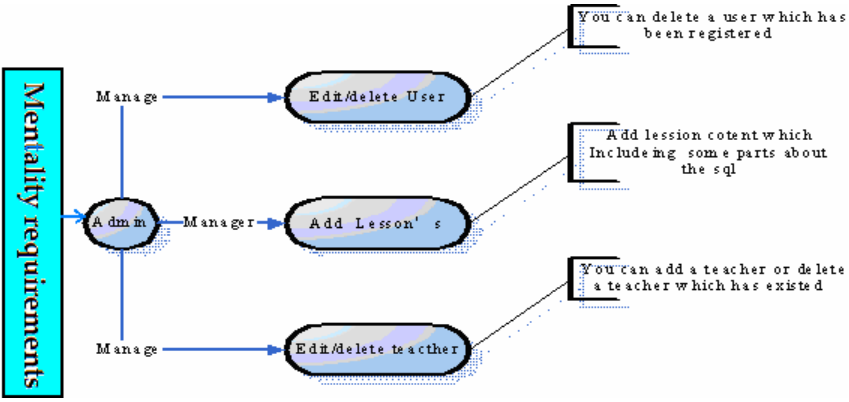


Fig. 5. The Flow Chart of User Information Module

includes masking sensitive information such a subject's identity, no capability to modify system data, etc [12].

Internally e-learning system is organized into several subsystems or modules, the most important being E-learning module and User information module. Since our system is build with the open architecture and different databases can be used as a storage medium. Our current implementation uses MySQL [13], an open source database, which offers substantial power to handle large amounts of data in an efficient manner.

Table 4. The Table of Records

Table	T_Record			
Column	Description	Type defination	null/not null	Restric-tion
userid	User id	int	not null	
Lesson_id	Lesson id	int	not null	

键	ID	名称	数据类型	大小...	空	默认值
		userid	int	4	✓	
		lesson_id	int	4	✓	

5 Major Function—Pronunciation Evaluation

In our humanized system, we have developed a teacher agent to fulfill pronunciation assessment as the main feature of our pervasive computing Mandarin e-learning system. It is composed with two modules: speech analysis and tone evaluation. Each of them can give a real-time result of learner's pronunciation, which is the basis of the next learning.

Ⅰ. Speech Analysis

Our speaker-independent Mandarin speech recognition component is firmly based on the principle of statistical pattern recognition [14]. When students' utterance inputs into system, a front-end signal processor with a sequence of acoustic vectors converts the speech waveform and the language model computes its probability. For each phone there is a corresponding statistical model called a hidden Markov model (HMM). The sequences of HMMs needed to represent the postulated utterance are concatenated to form a signal composite model and the probability of that model generating the observed sequence is calculated.

Our design choice is between two commonly used recognition algorithms: Viterbi and N best [14]. The Viterbi algorithm is fast, straightforward, and yields the single most likely spoken sentence, given the observation sequence and the HMM. The core of the Viterbi algorithm is to recursively compute the state probabilities *P(Oi,s)*, where n(s) is the initial probability of states.

$$P(O_1, s) = \pi(s) * P(o_1 \mid s) \tag{1}$$

$$P(O_i, s) = MAX_{p \in pred}[P(O_{i-1}, p) * A(p, s)] * P(o_i \mid s) \tag{2}$$

Considering the result of viterbi decode, our mandarin e-Learning system can evaluate the speech recognizer result and provide the speech accuracy of pronunciation to learners by using the $P(O_i \mid s)$.

Ⅱ. Tone Evaluation

Mandarin is tonal language. The Mandarin four different tones include a lot of important information. So we process the tone recognition separately. The pitch contrail can distinguish four Mandarin tones efficiently, and then an event detection pitch detector based on the dyadic wavelet transform which can detect the catastrophe point of speech signal when people speak [15]. We use cured-fitting technology to realize the classification of the different tones.

Observing the pitch contrails, we find the slope of contrail can distinguish the different tones easily when we regard the information about the relative phoneme of tone as in Fig.6. So we decide to extract the slope of pitch contrails to recognize the Mandarin four tones.

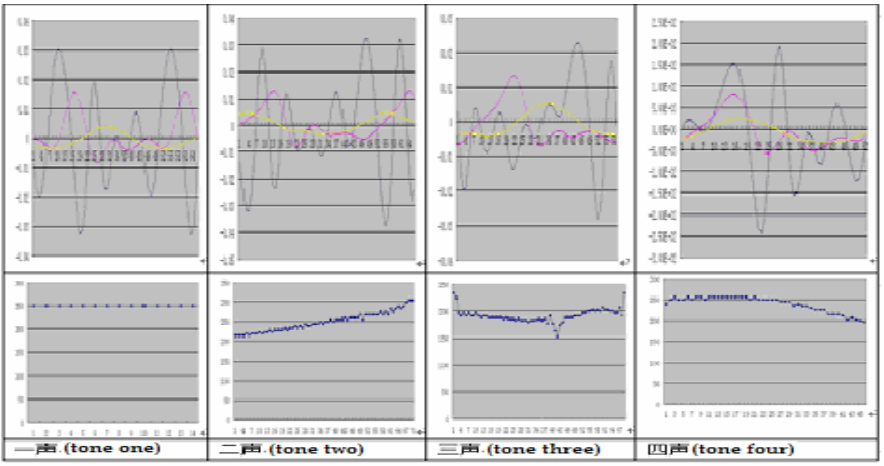


Fig. 6. The Pitch Contrail and Tone Results

Then we can calculate the tone accuracy of pronunciation by using the result of tone recognition. We set the average of these radian A_i is $\overline{A}$, the absolute value of $\overline{A}$ is $|\overline{A}|$.

$$|\overline{A}| = |(\sum_{i=1}^{n} A_i)/n| \tag{3}$$

When the tone is tone one, the accuracy of tone one is $|\overline{A}|/(\pi/18)\times 100\%$. When the tone is tone two, if $|\overline{A}| \geq \pi/3$, the accuracy of tone two is 100%, if $|\overline{A}| < \pi/3$, the accuracy of tone two is $(|\overline{A}| - \pi/18)/(\pi/3 - \pi/18)\times 100\%$. When the tone is tone four, the method of compute accuracy is the same as tone two. When the tone is tone three, the method of compute accuracy became complex. We need to find the lowest sample B, divide the A_i sequence into two sub-sequence

$A_i^1 = A_1, ..., B$ and $A_i^2 = B, ..., A_n$, then compute the $|\overline{A^1}|$ and $|\overline{A^2}|$. Set variable $R = \pi - |\overline{A^1}| - |\overline{A^2}|$, if $0 < R < \pi/3$, the accuracy of tone three is 100%, else the accuracy of tone three is $(R - \pi/3)/(\pi - \pi/3) \times 100\%$

6 Mandarin e-Learning System Design and Implementation

Based on the above design discussion, we design and implement our pervasive Computing Mandarin e-Learning System. When the learner enters the home page, he can start Mandarin learning depicted in the Fig.7. Fig.7 is an example of Mandarin e-Learning lessons. On the top of the page, it is the contents of the lesson, includes Chinese words, English words and bugles when the learner clicks them, he can hear natural pronunciation. On the bottom of the page, it is a text field, which is used to display the overall pronunciation of each utterance rated on a scale of 1-100. In terms of the related theories of speech analysis [14] and tone recognition [15], the e-Learning system will process the learner's Mandarin speech and the result.

Fig. 7. Learner's pronunciation Exercise Page

• Soft Phone: If the learner meets some difficulties during the course of language studying and clicks this button, a soft phone is started up via IP-to-PSTN gateway, and then he can communicate with a virtual teacher about his questions.

• Voice Mail: Java Mail used to design this button. Information of learners (time, space, the result of studying, etc.) can send to both the teacher and learners via email to make Mandarin learning conveniently and efficiently.

The Mandarin E-learning system also provides many smart servers. He can see his learning information, which is collected automatically for his teacher making the teacher know his learning effects. One important feature is that he can test his Mandarin pronunciation accuracy by the PRACTISE button as is shown in the Fig.7.

7 Conclusion

This paper describes a pervasive computing system design and implementation for Mandarin e-Learning. We analyze the foreign learners' requirements and discuss the implementation in pervasive environment. Learners will mainly use the platform for learning, because within the platform not only the learning material is presented, but also communication and interaction takes place. This gives us bright hope in the success of our scheme and we are convinced that such a scheme will indeed become practical and scalable for its deployment over Internet for Mandarin e-Learning. It's a place where foreign learners can find the solutions they are looking for.

References

1. http://www.go2china.net
2. http://www.cowob.com
3. http://www.learnchinese.cn/agents.html
4. http://www.chinesehorizon.com/
5. http://www.ottawa-mandarin-school.ca
6. http://www.easemandarin.com/
7. Loidl, S.: Towards pervasive learning: WeLearn.Mobile. A CPS package viewer for handhelds. Journal of Network and Computer Applications 29, 277–293 (2006)
8. Kourouthanassis, P.E., Giaglis, G.M., Vrechopoulos, A.P.: Enhancing user experience through pervasive information systems: The case of pervasive retailing. International Journal of Information Management 27(5), 319–335 (2007)
9. Su, C.J.: Mobile multi-agent based, distributed information platform (MADIP) for wide-area e-health monitoring. Computers in Industry, Corrected Proof (Available online September 17, 2007) (in press)
10. Schmidt, A., Terrenghi, L., Holleis, P.: Methods and guidelines for the design and development of domestic ubiquitous computing applications. Pervasive and Mobile Computing (Available online July 22, 2007) (in Press)
11. Ming, Y., Miao, Z.: Humanied Mandarin E-Learning Based on Pervasive Computing. In: I3E 2007: The IFIP Conference on e-Business/Commerce, e-Services, e-Society, Huazhong Normal University, Wuhan, China, October 10-12 (2007)
12. Thriskos, P., Zintaras, E., Germenis, A.: DHLAS: A web-based information systemfor statistical genertic analysis of HLA population data. Computer Methods and Programs in Biomedicine 85, 267–272 (2007)
13. Johnson, R.: Expert One-on-One J2EE Design and Development, Wrox-2003 (Chapter 9: Practical Data Access)
14. Young, S.: Large Vocabulary Continuous Speech Recognition: a Review, Cambridge University Engineering Department, pp. 1–23
15. Ming, Y., Miao, Z., Su, W.: Tone Analysis for a Pervasive Computing Mandarin e-Learning System. In: The Second International Symposium on Pervasive Computing and Applications, Birmingham, UK (July 2007)

An Interactive Simulator for Information Communication Models

Mohamed Hamada

Languages Processing Lab, The University of Aizu, Aizuwakamatsu, Fukushima, Japan
hamada@u-aizu.ac.jp

Abstract. Information theory is the science which deals with the concept 'information', its measurement and its applications. In common practice information is used in terms of a Communication Model in which the emphasis lies on the transport of information, as generated by a source, to a destination. The communication system should transmit the information generated by the source to the destination as fast and accurately as possible. To achieve this goal several coding techniques were developed based on mathematical concepts. Due to this mathematical nature, information theory course is used to be taught by a traditional lecture-driven style. Studies showed that lecture-driven style is not effective with computer engineering students due to their active learning preferences. In this paper we introduce an interactive communication model simulator to facilitate teaching and learning of the basic concepts of information theory course. We also show the effectiveness of using the simulator in classroom.

1 Introduction

Information theory is the science which deals with the concept 'information', its measurement and its applications. In common practice information is used in terms of a *Communication Model* in which the emphasis lies on the transport of information, as generated by a source, to a destination. It addresses questions such as: How to transmit and store information as compactly as possible? What is the maximum quantity of information that can be transmitted through a channel? How can security be arranged? Etcetera [1]. In answers for these questions several coding algorithms (such as Huffman code, Shannon code, Fano code, etc.) and several concepts (such as the concept of information entropy, and the concept of channel capacity) were developed.

Information theory concepts are abstract in nature and hence used to be taught by a traditional lecture-driven style which is suitable for learners with reflective preferences. Since computer engineering learners tend to have strong active preferences (Rosati [14]), a lecture-driven teaching style is less motivating for them.

Our communication model simulator (CMS) is designed to tackle this issue and meet the active learning preferences for computer engineering learners. CMS can be used as a supporting web-based tool for active learning not only for information theory course, but also for several other courses such as courses in telecommunications, error

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 88–98, 2008.

correcting codes, image processing, and other related fields. Such courses cover a variety of topics including coding techniques, communication channels, information source, error-detection and correction, information entropy, mutual information, in addition to basic concepts of probability. We cover such important topics in our CMS environment. CMS is written in Java as an applet using Java2D technology of Sun Microsystems [6]. This implies that our CMS environment is portable, machine independent and web-based enabled which makes it useful tool as interactive learning environment for CS and CE learners.

In designing our CMS learning tools we considered the active construction learning model [3, 16] that has some basic design principles including the following.

1. Teachers act as facilitators not as knowledge transmitters. This means knowledge must be actively constructed by learners, not passively transmitted by teachers.
2. Assessment procedures should be embedded in the learning process and should consider learners' individual orientations.

To show the effectiveness of our CMS environment as a model of interactive learning tool, several classroom experiments were carried out. The preliminary results of these experiments showed that using our tools not only improved the learners' performance but also improved their motivation to actively participate in the learning process of the related subjects and think beyond the scope of the class.

The paper is organized as follows. Following this introduction, section two briefly explains the topics covered in our CMS including Huffman, Shannon, Fano and Arithmetic coding techniques, channel capacity, information source, and the basic communication model. In section three we introduce the communication model simulator. The performance evaluation of the simulator will be given in section four. Finally, we conclude the paper and discuss the results and possible future extensions in section five.

2 Information Theory

Information theory is the science which deals with the concept 'information', its measurement and its applications. Its main goal is how information can be transmitted, from source to destination, as reliable as possible. The basic communication model, shown in Fig. 1, has two basic components; *transmitter* and *receiver*. The transmitter is responsible for the formatting and sending of the data from the source through the channel. Then the receiver gets and re-formats the data to its original form and passes it to the final destination.

The transmitter component has several units. The *data reduction* unit is responsible for removing unnecessary data from the source. The *source encoding* unit formats the reduced data to become as compact as possible for faster transmission. To prevent improper use of the transmitted data, the *encipherment* unit is used. In practice, the data channel is subject to external noise which may cause data distortion, to help detect these data errors, the *channel encoding* unit is used. The receiver component has units with opposite functions to the transmitter component. The *channel decoding* unit uses the channel encoding information to detect and correct any possible errors.

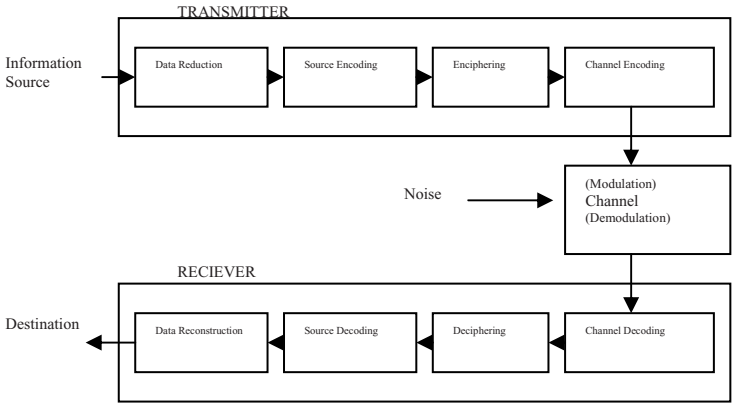


Fig. 1. The basic communication model

The encrypted data is then decrypted by the *decipherment* unit. To return the compacted data into its original format, the *source decoding* unit is used. Finally, the *data reconstruction* unit puts the data into its final form that is suitable for the destination.

Coding techniques are essential for reliable and efficient data transmission. Our CMS covers the basic coding techniques such as: Huffman, Fano, Shannon, Ziv-Lempel, and Arithmetic algorithms. Since we consider the noisy channel (which is more common in practice), the transmitted data are subject to distortion. To deal with this distortion we integrated the Hamming error-detection and correction techniques into our CMS environment.

Entropy is one of the most deep and fascinating concepts in mathematics. It was first introduced as a measure of disorder in physical systems but for information theory it will be most important in a dual role as representing average information and degree of uncertainty. We cover the concept of entropy and related concepts such as channel capacity and mutual information. Due to the lack of space we can't explain more about these concepts here. For more details we refer to [9].

3 Communication Model Simulator (CMS)

In this section we illustrate the communication model simulator (CMS). It simulates the operations of the basic communication model that is explained in section 2. Fig. 2 shows the user interface of the CMS.

To start using the CMS component, learners can click on the "Ctrl Panel" button, then the ctrl panel window appears (Fig. 3), from which they can start their first step of information theory learning. At first they must select a coding technique and then click on the "Set up" button. When the set up window appears (Fig. 4) they can input the source alphabet and its associated probability distribution.

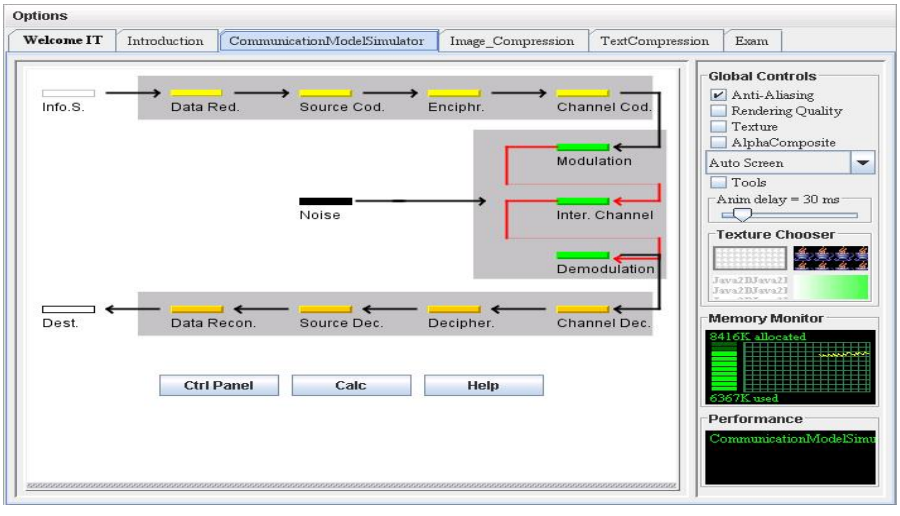


Fig. 2. The communication model simulator interface

Fig. 3. The control panel interface in the CMS

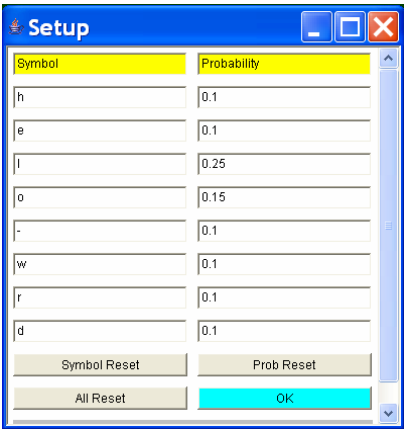


Fig. 4. The set up interface for the CMS

By clicking the “Message” button in the ctrl panel, learners can input the desired source message they want to send (Fig. 5). The “Send/Stop” button in the ctrl panel is used to start the transmission process of the inputted message. The transmission process is then visually seen on the CMS window in a step-by-step manner where users can pause and resume the transmission at any time to see which change had happened to the message at each point of the transmission.

Fig. 5. The input message interface for the CMS

The “Show” menu in the ctrl panel is used to view the details of many concepts such as the message transmission details (Fig. 6) and the coding/decoding processes details (Figs. 7 and 8). In addition to other concepts such as entropy, mutual information, channel capacity, and so on.

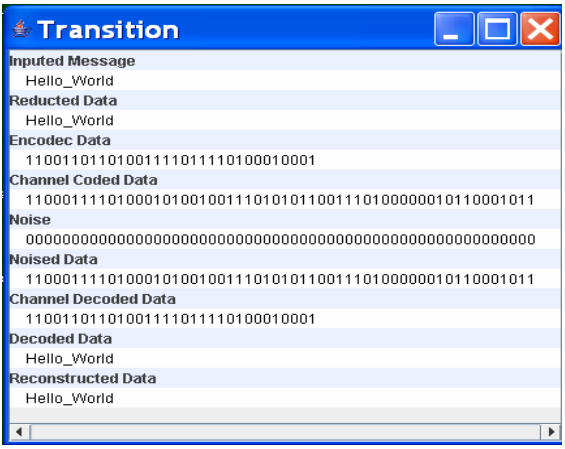


Fig. 6. The message transmission details

Figs. 4 to 8 show a detailed example. In this example the inputted source alphabet is h, e, l, o, - (for space), w, r, and d with probability distribution 0.1, 0.1, 0.25, 0.15, 0.1, 0.1, 0.1, and 0.1 respectively (Fig. 4). The inputted source message is “Hello World” (Fig. 5). The transmission details of this message are shown in Fig. 6. During the transmission process of this message the Huffman coding technique was selected by the learner. Figs. 7 and 8 show the details of the Huffman coding/decoding processes (of the source alphabet h, e, l, o, -, w, r, d) respectively.

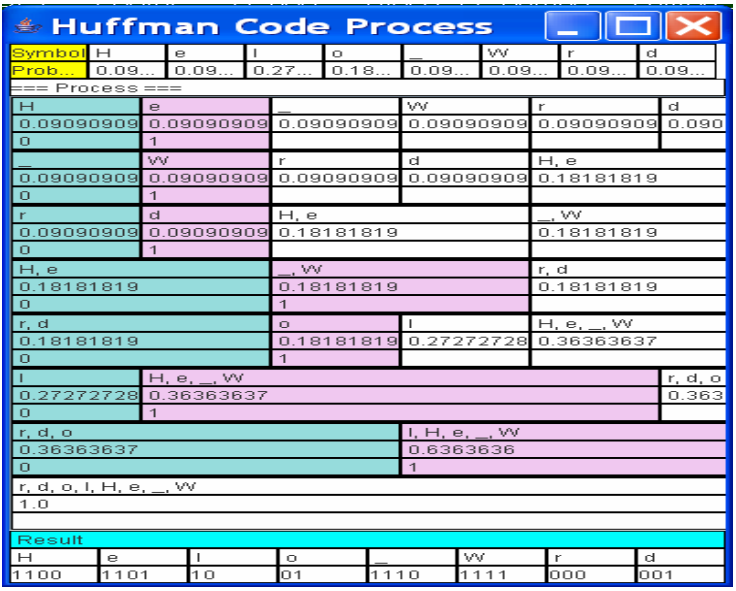


Fig. 7. The coding process details

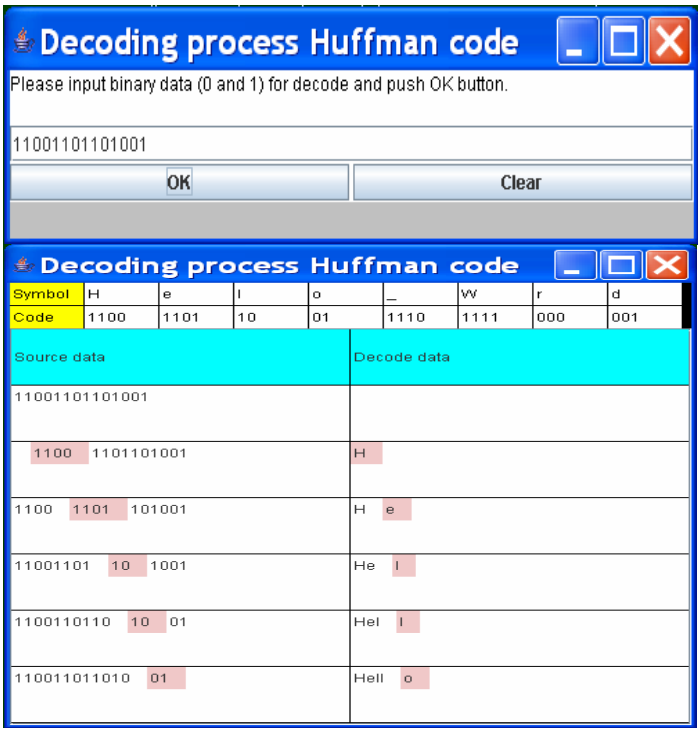


Fig. 8. The decoding process details

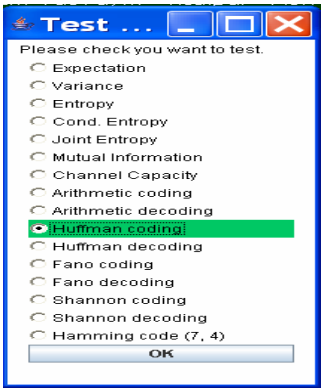


Fig. 9. The test interface of the CMS

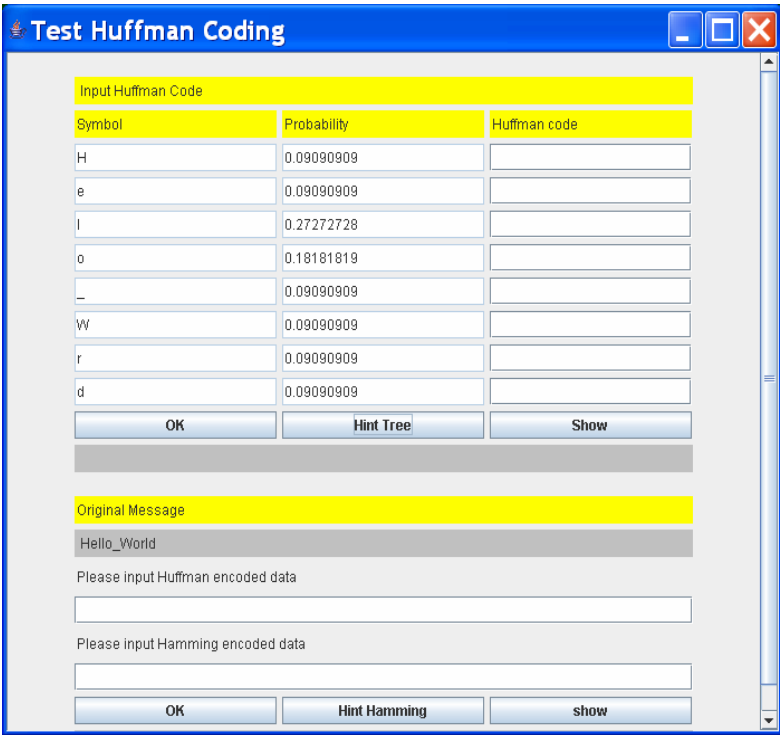


Fig. 10. The coding test interface (Huffman code example)

In addition to viewing the details, learners can try the concepts themselves through the interactive "Test" tool. From the ctrl panel interface learners can click on the "Test" button to view the test menu (Fig. 9). From the test menu learners can select a topic and start the test. For example, in Fig. 9, the Huffman coding test was selected.

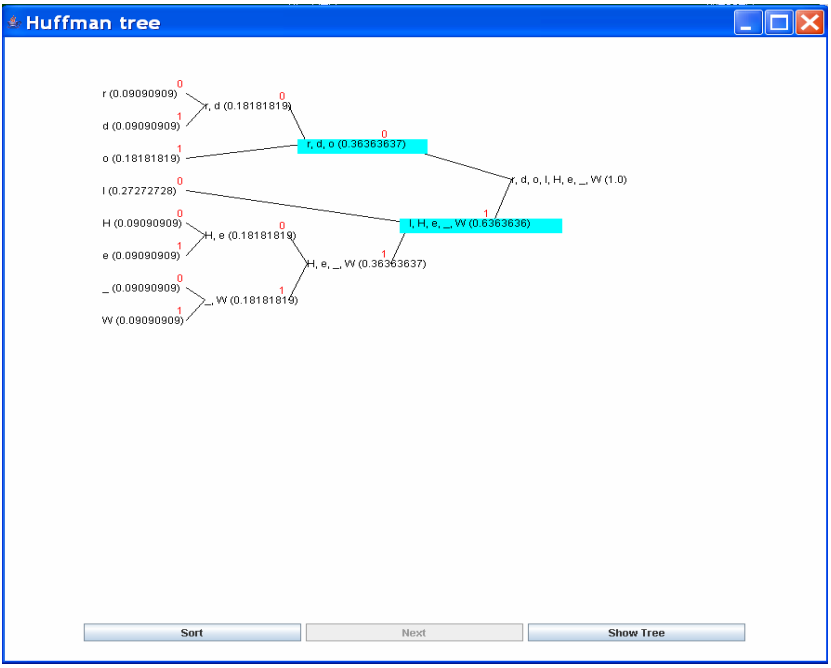


Fig. 11. The coding test detailed steps (Huffman code example)

From the Huffman coding test interfaces (Figs. 10 and 11) learners can try to find the correct Huffman code for the inputted source alphabet (h, e, l, o, -, w, r, d). They can repeat the trail (with some possible hints) until finding the correct code. Fig. 11 shows an animated binary tree that represents the Huffman code. Learners can also try to construct the tree to find the correct code.

4 Evaluation

We carried out two experiments and an opinion poll in order to evaluate the effectiveness of our CMS tools on the learning process of engineering students. The first experiment evaluates the learning preferences of the students according to the "Learning Style Index" of Felder-Soloman [15]. The second experiment evaluates the effectiveness of using the tools on the students' performance. Finally an opinion poll was carried out in order to observe the students' feedback about the tools.

To help learners find their learning preferences, the Felder-Soloman Index of Learning Style [15] was introduced. Fig. 12 shows a summary of the learning style quiz results from the author's evaluation as well as the data found at the University of Western Ontario by Rosati [14] where he surveyed 800 students and found that engineering students typically have preferences toward active, visual, and sensing learning preferences. It also contains a similar data by J. Masters *et. al.* at San Jose University [14]. It is clear that our data (on Japanese students) and Masters's data (on American students) support the data collected on Canadian students by Rosati [14].

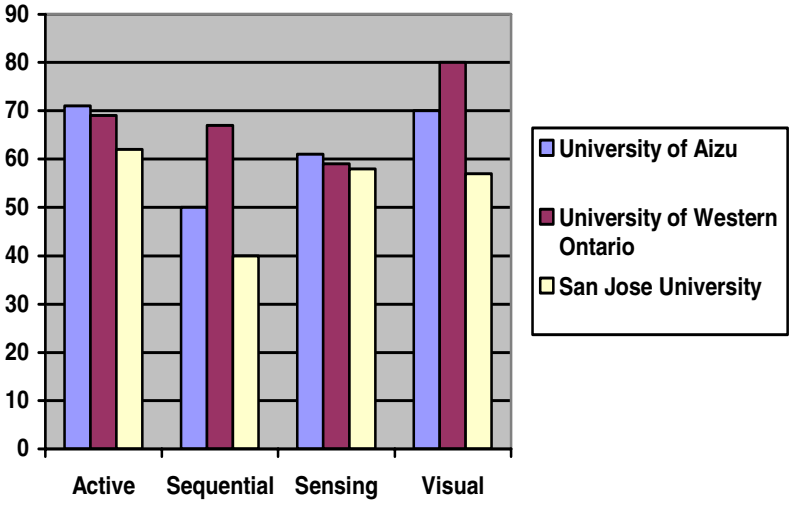


Fig. 12. Learning preferences

A preliminary study shows that using the CMS can improve the learning process of computer engineering students who study the information theory course and related courses. Last semester, 100 students in the information theory class were divided into four groups, each group containing 25 students. A set of 40 randomly selected exercises was distributed among the groups, 10 for each group. Each group members could collaborate inside their group but not with any other group members. No group could see the exercises of other group. Two groups were asked to answer their assigned exercises using the CMS and the other two groups without using it. An equal time period was provided to all the groups. The result of the answers showed a better performance for the two groups using the CMS. Then, the experiment was repeated by redistributing the exercises among the four groups. Again, the two groups with the CMS showed better performance in answering the assigned exercises.

In addition to the experiments an opinion poll among the learners was carried out. The result of the poll is shown in Table 1. Among the 100 learners in the class, 95 had completed the poll as shown in Table 1(a). Among the 95 responses 79% preferred using the CMS tools as shown in table 1(b). Most questions on the opinion poll were Likert-type questions that made a clearly negative or positive statement about the CMS tools and allowed the learners to strongly agree, agree, be uncertain, disagree, or strongly disagree. Scores for the CMS tools were generated based upon the learner responses. The scores could fall between 0 (worst) and 50 (best) and were divided into five ranges: greatly disliked the tools (score: 0-10), disliked the tools (score: 11-20), uncertain in preferences for the tools (score: 21-30), liked the tools (score: 31-40), and greatly liked the tools (score: 41-50). Table 1(c) shows the average score for the CMS tools lies on the far end of the "liked the tools" range.

Table 1. Results of the opinion poll

a.	Learners who completed the opinion poll	95 (out of 100)
b.	Learners who preferred the CMS tools	79%
c.	Average score for the CMS tools	40 (out of 50)
d.	The CMS tools made concepts easier to understand	85%
e.	The CMS tools made me think outside of the class	83%

Table 1(d) and (e) show the responses to other important questions. These results show that the majority of learners found that the CMS tools helped clarify important concepts and encouraged them to think about concepts outside of class. The latter is a significant accomplishment that could lead learners to seek more knowledge and information on their own.

5 Conclusion

With the vast advance in technology, the traditional lecture-driven classroom is giving way to a new and more active environment, where students have access to a variety of multimedia and interactive course materials. Such interactive course materials have already been introduced for several topics in engineering courses; see for example [2, 4, 5, 7, 8, 10, 11, 12, 13].

In this paper, we followed the same path and introduced a communication model simulator to support active learning in the information theory course. It can also be used in other related courses such as telecommunication, error correcting codes, image processing, and other similar courses. Our CMS environment is web-based, easy-to-use, and stand-alone which make it a useful tool of e-learning. Through the results of our experiments, we also showed that our CMS tools can enhance learners' motivation and performance. In addition an opinion poll showed a positive feedback on the CMS tools from the students. In future work, we plan to enhance our CMS tools by adding more features and more visual examples, and by performing more performance evaluation experiments.

References

1. der Lubbe, J.: Information Theory. Cambridge University Press, Cambridge (1997)
2. Hadjerrouit, S.: Learner-centered Web-based Instruction in Software Engineering. IEEE Transactions on Education 48(1), 99–104 (2005)
3. Hadjerrouit, S.: Toward a constructivist approach to e-learning in software engineering. In: Proc. E-Learn-World Conf. E-Learning Corporate, Government, Healthcare, Higher Education, Phoenix, AZ, pp. 507–514 (2003)
4. Hamada, M.: Visual Tools and Examples to Support Active E-Learning and Motivation with Performance Evaluation. In: Pan, Z., Aylett, R.S., Diener, H., Jin, X., Göbel, S., Li, L. (eds.) Edutainment 2006. LNCS, vol. 3942, pp. 147–155. Springer, Heidelberg (2006)
5. Head, E.: ASSIST: A Simple Simulator for State Transitions. Master Thesis. State Univesity of New York at Binghamton (1998), `http://www.cs.binghamton.edu/~software/`

6. Java2D of Sun Microsystems, http://www.sun.com
7. Java Team, Buena Vista University, http://sunsite.utk.edu/winners_circle/education/EDUHM01H/applet.html
8. Li, S., Challoo, R.: Restructuring an Electric Machinery course with Integrative approach and computer-assisted Teach Methodology. IEEE Transactions on Education 49(1), 16–28 (2006)
9. Mackay, D.: Information Theory, Inference, and Learning Algorithms. Cambridge University Press, Cambridge (2003)
10. Masters, J., Madhyastha, T.: Educational Applets for Active Learning in Properties of Electronic Materials. IEEE Transactions on Education 48(1) (2005)
11. Mohri, M., Pereria, F., Riley, M.: AT&T FSM Library. Software tools (2003), `http://www.research.att.com/sw/tools/fsm/`
12. Nelson, R., Shariful Islam, A.: Mes- A Web-based design tool for microwave engineering. IEEE Transactions on Education 49(1), 67–73 (2006)
13. Rodger, S.: Visual and Interactive tools. Website of Automata Theory tools at Duke University (2006), `http://www.cs.duke.edu/~rodger/tools/`
14. Rosati, P.: The learning preferences of engineering students from two perspectives. In: Proc. Frontiers in Education, Tempe, AZ, pp. 29–32 (1998)
15. Soloman, B., Felder, R.: Index of Learning Style Questionnaire, `http://www.engr.ncsu.edu/learningstyle/ilsweb.html`
16. Wilson, G. (ed.): Constructivist Learning Environments: Case Studies in Instructional Design. Educational Technology, Englewood Cliffs (1998)

iThaiSTAR – A Low Cost Humanoid Robot for Entertainment and Teaching Thai Dances

Chun Che Fung[1], Thitipong Nandhabiwat[2], and Arnold Depickere[3]

[1, 3] School of Information Technology, Murdoch University, Western Australia 6150
[2] Rangsit University, Pathumthani 12000, Thailand
l.fung@murdoch.edu.au, thitipong@rangsit.rsu.ac.th, a.depickere@murdoch.edu.au

Abstract. Development of humanoid and dance robots has improved greatly due to rapid advancement of electronics, computer, mechatronics and control technologies. While humanoid robots such as Honda ASIMO, Fujita HOAP-3 and Sony QRIO have dazed the public with their amazing capabilities, such robots are in very limited supply and they are also extremely expensive. On the other hand, the low cost toy robot, WowWee's Robosapien (RS), has become very popular. It has also expanded its functionalities in later models since its line of products were first launched in 2004. The most important aspect of such robot is its cost is only a fraction of the highly sophisticated robots. This study investigates the feasibility of using low cost robots such as RS for the purposes of entertainment and teaching Thai dances. Informal feedbacks and comments have shown a high degree of acceptance and keen interest. This demonstrates the potential of low cost robots for training, entertainment and edutainment purposes.

1 Introduction

Edutainment has long been recognized as a combination of entertainment and education. To most people, edutainment assumes the forms of TV programs, movies, video games and computer games. The educational aspects of edutainment programs are mostly based on "braingames" which aim to achieve certain intellectual learning objectives within specific curriculum. The learning outcomes of such "braingames" normally focus on academic goals such as improving literacy, numeracy and problem solving skills. On the other hand, physical education is another important aspect in the development of a student from early childhood to late adolescence [1]. Within a typical curriculum framework, dance plays a role in providing education and training in locomotion, balance, support, rotation, social and team skills. In the context of Thailand, traditional Thai dances form an important part of the national and cultural heritage which should be taught and preserved.

Development of dance robots has improved greatly due to rapid advancement of electronics, computer, mechatronics and control technologies. However, most humanoid robots such as Honda ASIMO, Fujita HOAP-3 and Sony QRIO are extremely expensive and are beyond the reach of the public or educational

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 99–106, 2008.

organizations such as schools. On the other hand, low cost toy humanoid robots have been gradually gaining acceptance in the consumer markets. In addition, such robots have also been expanding their functionalities with each new version being produced. The most important fact is that these robots cost only a fraction of the more advanced research humanoid robots. It is therefore hypothesized in this study that low cost robots could be used as an edutainment tool to motivate and encourage school children to participate in the dance sessions of the physical education curriculum. It is proposed that this category of robots could form an integral tool for entertainment and for the teaching of Thai dances in Thailand.

In this paper, a low cost toy humanoid robot, RS Media [2], is used to teach Thai dances and to entertain in public exhibitions. In order to establish a unique identity and character, the robot is dressed in traditional Thai custom and was given the name "iThaiSTAR". The name stands for "intelligent Thai Sanook Training-Assist Robot". Sanook in Thai means fun and entertaining. iThaiSTAR has performed in schools and danced with primary school children. It has also taken part in a major Exhibition at Bangkok – the ICT Expo [3], and was invited to appear at TV shows. It has also been reported in articles on the Web and in printed newspapers. Informal feedbacks and comments from the public mainly concerned the acceptance of such robots and the feasibility of using them for the purpose of teaching and entertainment. A high degree of acceptance and keen interest from the public was demonstrated. This could lead to the potential of using low cost robots for training, entertainment and edutainment purposes in the future.

2 Dance Robots and Robosapien

The ideas of using robots for entertainment and services are not new. Robot characters have been created and popularized in books, movies and TV. While these imaginary robot characters could do a lot in the movies or TV shows, their capabilities and functions are far more limited in reality. Nevertheless, there are steady improvement and progress in robot technologies aiming to develop the ultimate service or companion robot.

On the other hand, entertainment robots could be considered as one that entertains people and not just "as seen on TV". An example is "*Keepon*" which was developed aiming to provide dance-oriented nonverbal play between "itself" and children [4]. Their study demonstrated that the rhythmic synchrony of the robot's movements with the music and the children would have an effect on the quality of interactions and the rhythmic behaviors of children. Furthermore, the analysis and observations suggested that music provided a powerful environmental cue for the negotiation of rhythmic behavior relative to that of the robot; and that the robot's responsiveness to people's behaviors positively affected their engagement with the robot [5, 6]. On this basis, dance and music could be considered as a conducive means to facilitate robot and human interaction. This will serve the purpose of education and entertainment. Conducting research on dance robots certainly introduces the development of a new type of communication between human and robots. The role of robots in dance entertainment allows human to become both entertainers and spectators. Human behaves as a spectator when watching a robot dances with its own autonomous

movements and interactive capabilities. This can be classified as a form of real-time entertainment. The designer of the robot could be regarded as an entertainer when the robot performs built-in pre-programmed sequence of dance motions. This is a form of non-real-time entertainment. A possible scenario would where an interactive dance robot in real-time entertainment could change its dance or response according to audience requests. Alternatively, it could sense the audience's mood and adapts its dancing behaviors to reflect the sensor inputs. An ideal model of a dance robot would be the one that could provide flexible entertainment that ranges between real-time and non-real-time entertainment [7].

Apart from Keepon, another example is the "*Hip Hop Dance Robot*" at Kwansei Gakuin University. It is a humanoid robot that is capable to display various dance performances by concatenating a set of different short dance motions called "dance units" [8]. Another example is at Tohoku University, a dance partner robot referred to as "*MS DanceR*" (Mobile Smart Dance Robot) has been developed as a platform for realizing effective human-robot coordination with physical interaction. [9]. The dance robot "*HRP-2 Promet*", developed by Tokyo University, can perform a traditional Japanese dance by captured human dance movements using video-capture techniques then convert the input into a sequence of robotic limb movements and fed into its processors [10]. Moreover, Tanaka et al. [11] created non-interactive and interactive (posture mirroring) dance modes for a Sony QRIO robot in a playroom with children.

All the robots mentioned above are either experimental research robots which are in limited supply or they carry very high price tags. In this study, we aim to examine the feasibility of using low cost robots for the purpose of implementing a dance robot for entertaining and training. Details of the robot are described in the following section.

2.1 WowWee and Robosapien

WowWee Limited is a privately owned company and it is best known for its line of Robo-branded consumer robotic products. WowWee's Robosapien was the first line of robotic products that use biomorphic motion technology and is programmable to perform a variety of functions. The latest Robosapien models are the V2 and RS Media. Over 50 million RS units have been sold since its release in 2004. The Robosapien V2 is larger robot with an expanded list of English verbalizations. RS V2 also introduced basic color recognition sensors, grip sensors, sonic sensors and a wider variety of possible movements [12]. The RS Media has a body very similar to RS V2, but an entirely new processing unit based on a Linux kernel. RS Media is equipped to be a media center and has the ability to record and playback audio, pictures and video.

While many mass-produced humanoid robot kits or humanoid robots developed so far are inarguably flexible and capable to perform movements that are human alike, many of them are small in size, or do not possess the ability to produce audio independently. As it is intended in this study that the robot used in this research should be accessible by the public and be able to perform the Thai dance autonomously without the need of a computer or external speakers. Thus, this makes a robot selection process currently limited to Robosapien series, in particular the RS Media. In 2007, a special Robot Extension SDK written by Sun Microsystems was

released and it was bundled with the RS Media. 200 were available for sale at the Java One conference. In order to take the full programming capability of the robot, the RS Media Java SDK was acquired for this research.

2.2 Robosapien Media Features and Specifications

RS Media has a total of 11 degrees of freedom with measurement of approximately 58 cm. in height and 5 kg in weight. The low center of mass makes RS Media very stable. It has a Linux operating system with two 32-bit processors for handling the control of sensors and movements. RS Media has a vision system with a built-in full-color camera into its chest and face-tracking intelligence. It can play MP3 music through its multiple speakers and back-mounted woofer as well as displaying photos and MP3 information. The user may play Java games and MP4 video on its 1.9-inch 16-bit color LCD screen. It also has 40 MB of internal flash memory with the ability to utilize the storage in a 1 GB SD card in its external card slot.

RS Media comes with three distinguish sensors: sight, sound and touch. The motion and color tracking, and, sound localizing sensors are unique features for interactivity communication with the user. They could incorporated in the dance sequence for changing of dance styles. It also has an infrared vision system which could be used to differentiate between certain colors. When the robot is stationary, the infrared system can detect movement at two different ranges. The sensors in his feet will also detect objects that the robot has encountered and it will stop. Sensors are also built in to the hands so that it knows when it's picked it up successfully, if not, the robot will provide an audio feedback. According to RoboGuide [13], the RS Media also has a range of varied internal sensors, including 4 pots, 3 tilt switches, 5 encoders and 2 switches. The main movements are: Neck, Shoulder, Wrist, Hands, Waist and Foot. The RS Media also has a range of external sensors, including 8 touch sensors (button), 3 sound sensors (microphone) and 3 sight sensors (IR Receiver) located in its Head, Chest, Hand and Foot.

It can be concluded that the RS Media provides a fair degree of flexibility for control and monitoring its movement. In addition to its multimedia capabilities, this makes it suitable for the investigation in this project.

3 Thai Dance and Limitations of RS Media

3.1 Historic Background of Thai Dances

Thai dance, or "*Ram Thai*" in Thai language, is the main dramatic art form of Thailand and it is considered as one of countless worldwide dance types in existence. Thai dance on its own, likes many forms of traditional Asian dance, can be divided into two major categories that correspond to the "high art" or classical dance, and, "low art" or folk dance [14]. Thai classical dance includes main dance forms like "*Fawn Thai*" accompanied by folk music and it varies according to the style of the particular region. "*Khon*" is the most stylish form of Thai dance performed by group of non-speaking dancers while a story is being told by a chorus at the side of the stage. "*Lakhon*" has costumes which are identical to Khon, but Lakhon dance

movements are more graceful, sensual, and fluid. The upper torso and hands are particularly expressive with movements portraying specific emotions; etc. Thai folk-dance includes main dance forms such as "*Ram*", which are originated from numerous regional dances. "*Likay*", contains elements of pantomime, comic folk opera, and social satire. They are generally performed against a simply painted backdrop during temple fairs. "*Ram Muay*" is the ritualized dance that takes place before Southeast Asian kickboxing matches such as Muay Thai. "*Wai Khru*" is a ritualized form of dance meant to pay respect to, or homage to the "*khru*" or teacher. It is performed annually by Thai classical dance institutions; etc.

As there are many varieties of Thai dances, it is vital that a careful selection of the one to be used in this research is properly consulted with a Thai dance professional. Thai dance ranges from a simple movement to complex movements that if an inappropriate one is chosen, the robot might not be able to perform due to its physical limitations.

3.2 Implementation of Thai Dance by iThaiSTAR

After a careful study, *iThaiSTAR* is currently capable to demonstrate the Thai folk-dance performance called "*Ram Wong*". The art of *Ram Wong* is originally adapted from "*Ram Tone*", where it uniquely specifies that dancers must follow the rhythm of the tone drum which is especially made for the dance. *Ram Wong* is one of Thailand's most popular folk dances. It has been popular among Thai people in some regions of Thailand. This Thai dance used to be played with the performance of Thai traditional music instruments consists of "*Ching*" (a kind of Thai important percussion instrument made of metal), "*Krab*" (a kind of Thai important percussion instrument made of wood), and "*Tone*" (Thai drum made of carved wood or baked clay). In 1940, *Ram Wong* dance pattern influence spreads to the other regions in Thailand. It has effectively become very popular among the people in every region of the country and has created the rhythmic dialogue to sing together with the performance of Thai music. Basically, the dialogues are about persuasion, teasing, praising, and parting. *Ram Wong* has been very popular among the people in central region of Thailand during the World War 2 (1941-1945). Owing to the support of the government, *Ram Wong* has been reformed by the Fine Arts Department of Thailand in 1944. At that time four new rhythmic dialogues had been created. The songs and music instrument had been adapted to be more contemporary. Some movements such as "*Tar Sod Soi Ma La*", "*Tar Ram Sai*", etc., had been settled as standard patterns of *Ram Wong*. The name *Ram Tone* (the tone dance) had been changed into *Rum Wong* (the circle dance) because of its movement which people was often dancing around like making a circular movement. Later, Premier Piboonsongkram created six more new rhythmic dialogues introducing the *Ram Wong* as a modern Thai dance. Finally, the *Ram Wong* has ten songs with specific movement patterns where the dancers moving round in a circle. The song lyrics refer to the goodness of Thai culture and the ability and daring of Thai warriors. After World War 2, *Ram Wong* has been kept active among people until now. *Ram Wong* is widely performed not only by Thai people, but also foreigners in dancing ball. Many *Ram Wong* songs have been created by following the ten specific standard movement patterns.

The two *Ram Wong* songs that *iThaiSTAR* could perform, each acquired from the ten specific preserved standard movement patterns. The two songs are "*Ngam Sang Duen*", which obtained the standard movement called "*Tar Sod Soi Ma La*" (imitated from the actual actions of the local people making a flower garland by having one of the hands holding a cotton string and another hand pulling a flower on the string outwards from the body towards the side). The song "*Ram Si Ma Ram*" synchronizes with the standard movement called "*Tar Ram Sai*" (imitated from the actual actions of people trying to persuade one another by stretching both arms almost parallel to the ground and twisting both hands up and down opposite one another) This is shown in Fig. 1.

Fig. 1. Main movements of "*Tar Sod Soi Ma La* [15]" and "*Tar Ram Sai* [16]"

3.3 Limitations of RS Media

Even though Robosapien Media's humanoid body movements such as bending, sitting, standing, lying down, getting up, dancing and waving are improved in many ways over the previous versions and other similar low cost robots, there are many still limitations. These constraints are described as followings.

The RS Media only has a total of 11 DOF, of which each leg only has 1 DOF. This is much less in comparison to ASIMO, QRIO and HOAP-3. These three robots have 34, 38 and 28 DOF respectively. Hence, many of RS Media's movements especially leg movements are impossible. For example, it cannot turn left or right the same way as a human, therefore, it would have to walk backwards and twist the waist at the same time in order for it to turn. Moreover, the RS Media's ability to walk is still limited such that it waddles excessively. In addition, the surface on which the RS Media walks on has an effect towards its leg movements as it cannot dance effectively on the "tend-to-sticky" kind of floors such as carpet.

The servos that are used offer quite a large degree of flexibility with the ability to grip objects, move its head, lean forward, move sideways and back, as well as wave its arms. The issues start to arise when the RS Media tries to pick something up. The object that the robot needs to pick up has to be positioned and shaped properly, otherwise the robot will struggle to pick it up or will not pick it up at all. Another problem is the energy consumption. If too many servos are running concurrently, this will tend to reduce the running time of the robot. While AC/DC adaptor could be

used, the servos for the legs will not be turned on thereby making the robot could only move its hands, head and upper body.

4 Feedbacks and Discussions

Initially, assessment of the performance of iThaiSTAR was based by comments from the dance professionals. Informal feedbacks and comments from the public were subsequently observed during public exhibitions as the audience approached the exhibitor for further information. The feedback collected cannot be considered as a comprehensive or objective assessment of the use of iThaiSTAR. It however gives an indication the feasibility of using low cost robots as hypothesized in this proposal. The general feedback provided an indication of the public's interest in the project. A summary is given below:

- 60% of those who gave comments are under 25 years old.
- Approximately 60% of those who provided feedback are female.
- Three quarter of the comments are seeing a dance robot for the first time.
- Over 80% of the comments considered iThaiSTAR is "Good" to "Great"
- Nearly 98% considered iThaiSTAR as "cute"
- Over 80% considered the dance is "Good" to "Great"
- Over half indicated that they will dance with the robot
- 80% commented they prefer the robot to a dog.
- Over 90% would like to have a robot in the house.

The answers from the public have clearly indicated a positive impression and general acceptance of the robot in their lives. It is also interesting to see the percentage of female as compare to the male in a ratio of 2 to 1. Also, it is surprising to see close to 80% of the comments preferred the robot to a dog! The cuteness and attractiveness of iThaiSTAR also received overwhelming favorable responses. They also perceive the robot could be useful for other services in daily life.

5 Conclusion

This paper has reported the initial phase of the development of iThaiSTAR, a low cost humanoid robot for the purpose of entertainment and teaching of Thai dances. The research work so far has been the investigation, design and development of using an off-the-shelf toy robot, WowWee's RS Media for the purpose of this study. The public appearances of iThaiSTAR have attracted much attention from the audience and the media at Thailand. Informal comments have been positive and expert comments on the dance performed by the robot are good. It is expected that the research will be embarked into its second phase with further development of content and objective assessment of the effectiveness of iThaiSTAR. Another related project, R4RE, is currently under development. The project will use similar robots for applied edutainment programs at rural communities in Thailand and Australia.

References

[1] Curriculum Council of Western Australia, Curriculum Guides and Elaborated Guides [Accessed: December 11, 2007], http://www.curriculum.wa.edu.au/pages/curric_guides/index.html
[2] WowWee [Accessed: December 11, 2007], http://www.wowwee.com/
[3] Bangkok ICT Expo [Accessed: December 11, 2007],
[4] http://www.bangkokictexpo.com/
[5] Michalowski, M.P., Sabanovic, S., Kozima, H.: A Dancing Robot for Rhythmic Social Interaction. In: Proceedings of the 2nd Annual Conference on Human-Robot Interaction (HRI 2007), pp. 89–96. ACM, New York (2007)
[6] LERN: Leading Edge Robotics News, Keepon Keeps Dancing On, Robot, issue 8, p. 15 (Fall 2007)
[7] Michalowski, M.P., Sabanovic, S., Michel, P.: Roillo: Creating a social robot for playrooms. In: Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2006), pp. 587–592. University of Hertfordshire, United Kingdom (2006)
[8] Shinozaki, K., Iwatani, A., Nakatsu, R.: Concept and Construction of a Dance Robot System. In: 2nd International Conference on Digital Interactive Media in Entertainment and Arts (DIMEA 2007), Perth, Australia (2007)
[9] Shinozaki, K., Iwatani, A., Nakatsu, R.: op. cit.
[10] Kosuge & Wang Laboratory, "MS DanceR," Research [Accessed: October 29, 2007], http://www.irs.mech.tohoku.ac.jp/research/RobotPartner/dance.html
[11] Randerson, J.: Japanese teach robot to dance. Guardian Unlimited (August 8, 2007) [Accessed: November 28, 2007],
[12] http://www.guardian.co.uk/technology/2007/aug/08/robots.japan
[13] Tanaka, F., Cicourel, A., Movellan, J.R.: Socialization between Toddlers and Robots at an Early Childhood Education Center. Proceedings of the National Academy of Sciences of the USA (PNAS) 104(46), 17954–17958 (2007)
[14] Hosein, K.J.: Review of the RS Media. RoboCommunity (January 20, 2007) [Accessed: September 22, 2007], http://www.robocommunity.com/article/10627/Review-of-the-RS-Media/
[15] RoboGuide, RS Media, RS Media RoboGuide (February 5, 2007) [Accessed: September 22, 2007], http://www.evosapien.com/robosapien-hack/nocturnal2/RSMedia
[16] Mahidol University, The Rituals and Traditions of Thai Classical Dance, Mahidol University [Accessed: September 23, 2007],
[17] http://www.mahidol.ac.th/thailand/classical-dance.html
[18] Banramthai.com, Ngam Sang Duen, Banramthai [Accessed: November 3, 2007], http://www.banramthai.com/html/rw_kham.html
[19] Banramthai.com, Ram Si Ma Ram, Banramthai. [Accessed: November 3, 2007],
[20] http://www.banramthai.com/html/rw_ramsi.html

The Study on Visualization Systems for Computer-Supported Collaborative Learning

SooHwan Kim[1], Hyeoncheol Kim[1], and SeonKwan Han[2]

[1] Dept. Comp. Edu., Korea Univ., Anam-dong Seongbuk-Gu, Seoul, Korea
[2] Dept. of Comp. Edu., Gyeong-in National Univ. of Education, Gyeyang-gu, Incheon, Korea
love0jx@hotmail.com, hkim@comedu.korea.ac.kr, han@gin.ac.kr

Abstract. In this study, we developed a visualization tool that shows learners' interactive activities graphically during computer-supported collaborative learning (CSCL). The tool helps teachers to monitor and evaluate the interaction activities among students, and thus it makes it easier to develop more effective instructional design for CSCL. For experiments, we used a web-based discussion group of 112 students. We investigated the effects of visualization of interactions on web discussion CSCL. Also, in lesson practice with this tool, we verified the effects through questionnaire and interviews. Our result show that the visualization tool provides a teacher with patterns of collaborative activities very clearly and insights of how to promote students' interactions more efficiently.

1 Introduction

Computer-supported collaboration learning (CSCL) concerns with how students can learn together with the help of computers and internets. CSCL tools can range from email, online discussion groups and Internet chat rooms to sophisticated group decision support systems. CSCL follows the socio-constructivist learning theory in that it brings learners together and offer creative activities of intellectual exploration and social interactions, and in that learning takes place naturally through social interactions among students. As CSCL develops, it becomes more apparent that a transformation of the whole concept of learning is required, including changes in instructional design, evaluation of learning activities, and roles of teachers and students [10]. CSCL stresses collaboration among the students, so that they are not simply reacting in isolation to posted materials. Thus, a teacher must motivate and guide students so that they learn by exploring questions and answers together, teaching each other and seeing how others are learning. However, it is not easy for a teacher to achieve stimulating and sustaining productive student interaction. There have been several CSCL tools, such as TeamWave Workplace, BSCW, Groove, Shadow Network Space, SCLIE, KIE and so on [4][9]. But they do not offer functions of instructional design, and thus it is difficult for teachers to evaluate and understand students' learning activities and to guide students during CSCL. COLER [3] and MASPLANG [6] provide some visualized features but they are very limited and hard to see interactions dynamically. To overcome the problem, we develop a visualization tool that shows the students'

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 107–113, 2008.

interactions graphically, so that a teacher can monitor and evaluate their collaborative learning activities in CSCL environment. For our experiments, we use a web-based discussion group board and the visualization tool to evaluate students' interaction activities. We also propose the learning steps that describe how to use the visualization tool effectively to promote interactions and feedbacks. Our experiment results show that the visualization tool is clearly useful for a teacher to understand students' collaborative interaction and thus learning activities.

2 Web-Based Discussion Learning for CSCL

There is no feedback in existing web-based discussion step in CSCL [5]. Therefore, we suggest web-based discussion step including feedback with the visualization tool. In the part of unfolding, until now the tools that an instructor can grasp interactions among learners at a glance are not offered. So an instructor can seize interactions simply by entering the chatting room and looking at the board. That is why not only an instructor cannot grasp real discussion's condition but also get information for the feedback. Once a discussion start, an instructor only use ways that enter the chatting room and view web-based discussion board as a promoter of interactions. Therefore in this step, instructors need a tool that can have glance interactions among learners and the learning step that give the feedback constantly. In this study, to solve these problems, we develop the visualization tool that shows interactions among learners and suggest the learning step applying this tool such as Table 1.

Table 1. Web-based CSCL step

Step	Essential step	Option step
preparation	Learning preparation Making teams	
introduc-tion	Grasping debate problem Choosing a position and Making small group Devising learning plan	
expand	Individual learning for debate Small group debating **Feedback through visualizing**	Learning from different angle for debate Debating from different angle
arrange-ment	Jointing result and arrangement Evaluating and reflection	

We included a step that an instructor grasps interactions among learners at a view and give the feedback, such as applauses about learners who are good at participating in the discussion and advices about learners who are not by applying the visualization tool during the process of the collaborative learning. Even though emphasize a role of an instructor who promotes participation of a discussion strategically, it is uncommon to show using a tool realistically in this study. In short, offering the visualization tool to

grasp interactions among learners is possible it. The feedback engages with learner's discussion activities, gives a proper reinforcement and promotes interactions among learners more briskly.

3 Design and Implementation

3.1 Overview of Systems

In this study, a system that we want to develop will follow a collaborative discussion learning model step. So we first decide to construct web site having a discussion board.

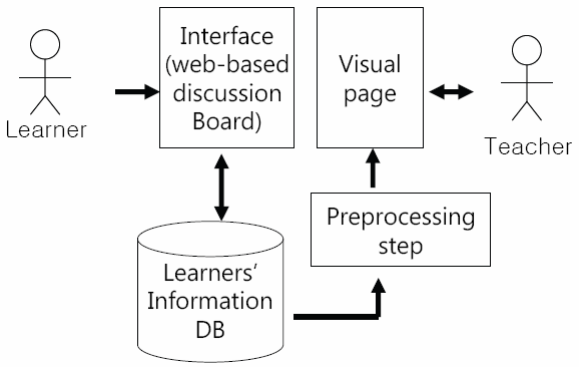
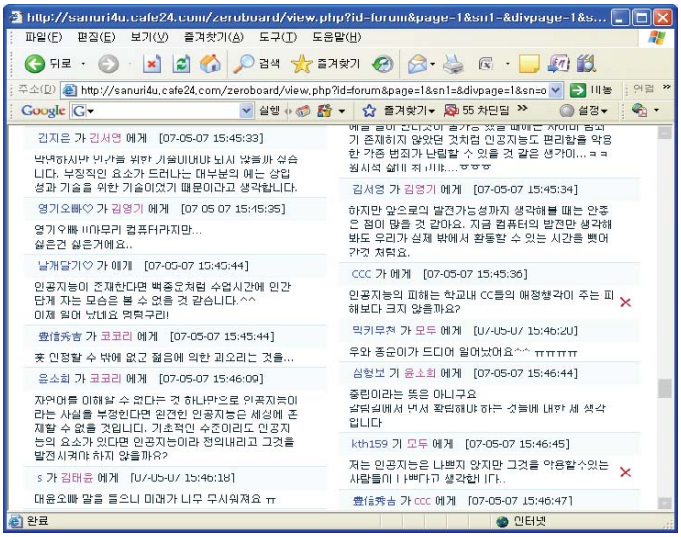


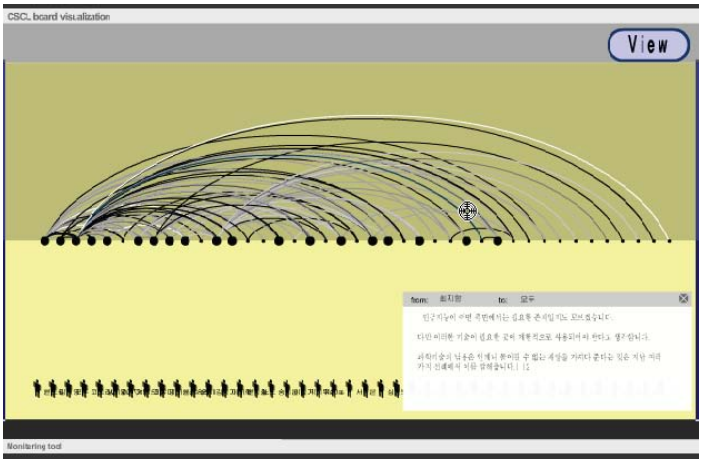
Fig. 1. Architecture of Systems

An instructor writes a subject of discussion on the web-discussion site, learners connect, click a certain subject and then enter web-discussion board. When they write an article on the board, they must appoint a receiver of discussion. Learner's name and conversation are stored in the learners' information DB. Then an instructor can develop a visual page based on information that can grasp discussion conditions at a look. Before the data passes a visual page, there is a handling step. In this step, if a learner writes curses or null message, they cannot write an article. It makes an instructor grasp meaning interactions. As form of debate discussion, a site for the CSCL was made to help a questioned person input and developed to let both a questioned person and an asking person appear. If a learner who participates in the discussion makes clear the sender and receiver, they can write an article. Information such as a person who writes an article, a receiver, a password, time to write and an agreement or opposition are stored in the learners' information DB. And it shows interactions depending on data, when an instructor pushes the button related in a visual page(shown as the right of Fig 2).

The A of Figure 2 is interface and shows the contents of discussion (Korean version). The subject of discussion is 'Would artificial intelligence lead our life to be comfortable?' In the B of figure 2, the left area is in approval and the right area is in opposition.



A



B

Fig. 2. The discussion board for CSCL(Korean version) and visual page

3.2 Implementation of Visualization Systems

The visualization page only will be offered to the instructors like as Figure 3, whenever they click view button. It will show relationship of interactions as visual form by bringing the learner's log information. Human's shape and names as shown in the bottom are learners who took part in the discussion. And we let the size of spot corresponded proportionately to amount of interactions. Because interactions among learners are described to a curve, learners can be connected between sender and

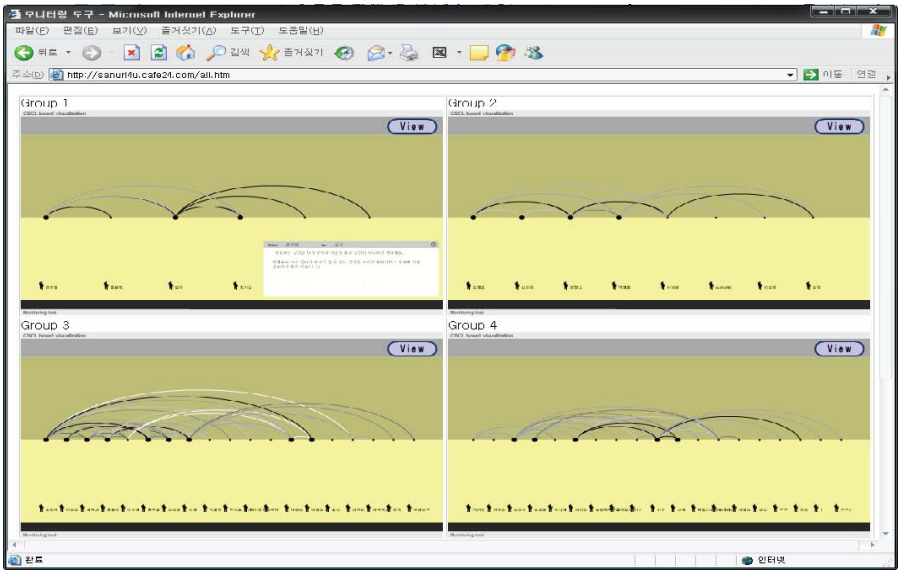


Fig. 3. A screen of monitoring tool

receiver. And the more amount of message, the more line color becomes dark. If an instructor put a mouse on the spot, he views learner's name. If they put a mouse on the curve, a message that a curve means appears in the part of right bottom. An instructor can give the feedback by grasping discussion active degree and conditions of participation.

4 Experiment

In this study, we progressed web discussion collaborative learning with 112 students in Gyeong-in National University of Education, who were juniors. The participants in subject discussion were made up 74 students 2 classes, them in debate discussion were made up 38 students of a class. The first subject of subject discussion is 'Would artificial intelligence lead our life to be comfortable?' The second subject is 'How to protect my pc safely?' Finally, the subject of debate discussion is 'In ubiquitous age, do we need school that is physical space?' when discussing, we gave 40 minutes. And in each class, there was a student who played teacher's role and were 3-4 assistant teachers. As analysis of visual tools, we focused on convenience, effectiveness and easy when using. Because this visual tool is given not to learners but to teachers according to learning steop, we surveyed and interviewed with 13 students who played teacher's role.

Figure 4 shows answers that 13users playing real teachers gave in an aspect of effectiveness, 84% of them gave positive answers. Also in an aspect of convenience 100%, and in that of convenient using method, 84% of them gave. Therefore, we realized that this monitoring tool was useful to observe learners and promote a discussion. Also, according to interview of users, they were thinking that this tool is effective, as follows.

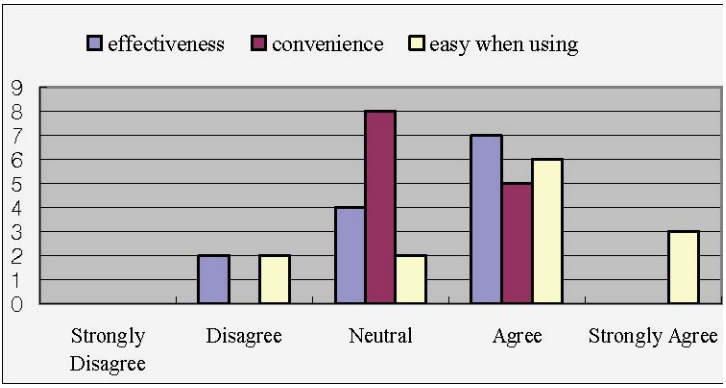


Fig. 4. Effectiveness of monitoring tool

- It was good that a teacher recognized whole discussion condition.
- By comparing, we could catch active group and non-active group in aspects of participation.
- We could recognize who took part in discussion intently.

Furthermore, they suggested several improvements about this tool, as follows.

- The message that causes an inactive learner to participate in a discussion actively needs to be transmitted directly on the monitoring screen.
- a restriction method is needed; "people distributing a discussion can't write down" etc.
- The orders of opinion exchange need to be showed.
- The method that users can express important opinion differently is needed

These points will help to supplement after system.

5 Conclusions and Future Works

In this study, we develop a visualization tool that grasps collaborative learning activities during web-based discussion sessions, and suggested a instruction-learning steps using the visualization tool. Also, by applying the developed tool to practical lesson, we verified convenience, effectiveness and easy when using through questionnaire and interviews. In this study, we experimented with 112 university students. The result clearly shows that this visualization tool is very useful to evaluate learners' interaction. In addition, user interview shows that this tool is effective, convenient and easy to use. Using this tool, a teacher can play a role of a guider and a supervisor of the collaborative learning. It can be used in any platform because it was based on the web. Also it can give a proper feedback to the student groups by analyzing patterns of learning activities. Finally, in the case of evaluate individuals or groups after CSCL, it provides a teacher with exchanged messages and grade of discussion participation

easily. As succeeding this study, we will study reformed improvements through interviews to improve this tool.

Acknowledgement

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOE) (KRF-2007-B00082).

References

1. Bastiaens, T.J., Martens, R.L.: Conditions for web-based learning with real events. In: Abbey, B. (ed.) Instructional and cognitive impacts of web-based education, pp. 1–31. Idea Group, Hershey (2000)
2. Davices, D.: Learning network design:coordinating group interactions in formal learning environments over time and distance. In: O'Malley, C. (ed.) Computer supported collaborative learning, pp. 101–124. Springer, London (1994)
3. Gauthier, G., Frasson, C., VanLehn, K.: In Intelligent Tutoring Systems. In: Proceedings of the 5th International Conference, pp. 325–333 (2000)
4. Jang, H.W., Suh, H.J., Moon, K.A.: Analysis of Collaborative Learning Model and Collaboration Tools in e-Learning Environment (2005)
5. Lee, I.-S., Leem, J.H., Jin, S., Sung, E.M., Moon, K.A., Seo, H.J.: Analysis of collaborative learning behaviors and the roles of collaborative-learning agent. In: Proceedings of e-Learn 2004. Association for the Advancement of Computing in Education, pp. 2748–2754 (2004)
6. Marzo, J.L., Peña, C.I., Aguilar, M., Palencia, X., Alemany, A., Vallès, M., Johè, A.: Adaptive multiagent system for a web-based tutoring environment - final report March (2003)
7. Lin, X., Hmelo, C., Kinzer, K., Secules, J.: Designing Technology to Support Reflection. Educational Technology Research & Development 47(3), 43–62 (1999)
8. Kim, M.-Y., et al.: New Flash Masters Of Korea 2005. The electronic Times (2005)
9. Gwon, S.-J., Kim, D.-S.: Development of Computer Supported Collaborative Learning Platform Prototype. Korean Association for Educational Information and Broadcasting 2001 7(1), 119–145 (2001)

Computer-Assisted Paper Wrapping with Visualization

Kenta Matsushima, Hiroshi Shimanuki, and Toyohide Watanabe

Department of Systems and Social Informatics,
Graduate School of Information Science, Nagoya University
{matsushima,simanuki,watanabe}@watanabe.ss.is.nagoya-u.ac.jp

Abstract. This paper proposes an approach to support a wrapping process from an object and a piece of wrapping papers by the computer animation method. The combinational information about an object and a piece of wrapping papers generates possibly multiple feasible wrapping processes. Therefore, we first argue the state of paper wrapping which represents each stage in the wrapping process, and then assume that sequences of ordered states, called state space, construct the wrapping process successively. Second, the folding operations called *wrap-folding* are defined. These operations generate multiple wrapping processes on a state space which are physically feasible. Third, an algorithm of searching effective wrapping processes which satisfy demand of user is described. Finally, we give some experimental results to make the practical efficacy of proposed method clear. With the processing flow, we describe a computer assisted paper wrapping proto-type system.

Keywords: Wrapping, process design, paper folding, design support.

1 Introduction

There are many ways to wrap not only box-types of goods but also various shapes of goods. The paper wrapping is performed by experts who have a lot of wrapping experiences generally. It is difficult for beginners to perform the paper wrapping because the paper wrapping is performed with various shapes of goods and a sheet of paper with the various sizes; additionally, the wrapping process is not unique and there are many wrapping processes. Therefore, the instruction books about paper wrapping are utilized for beginners as a tool for paper wrapping support. However, the instruction books do not show enough information about paper wrapping as assistant means for beginners: for example, only 2-D images about wrapping process may be shown or one of the examples is illustrated. In order to support paper wrapping activities, the method for representing paper wrapping processes effectively is necessary.

A lot of studies related to the paper folding, which is called Origami, have been conducted until now. Most of these studies are carried out by mathematicians: they attempt to elucidate the geometrical properties of Origami by use of mathematical methods [2,3]. The study which proposed the method for generating the folded Origami model from the information of crease lines on crease pattern has been conducted [5]. In this study, the folding processes are not so considered although the properties of crease lines are elucidated. As a study representing a folding process of Origami in 3-D virtual space, Miyazaki, et al. [4] developed a virtual interactive manipulation system for

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 114–125, 2008.

simple folding operations. A user folds a piece of papers by operating a mouse, and then the system constructs folded paper model. The user can get the objective model by determining all folding operations.

The most studies about Origami support are mainly utilized for Origami designers. In these studies, it is necessary to design folding operations, crease pattern and others exactly. On the other hand, in the paper wrapping, folding operation is constructed by stages, based on the information about the shape of target goods and the size of wrapping paper. Each folding operation depends on a condition of wrapping paper. Therefore, from a viewpoint of paper wrapping support, the method for managing these wrapping processes effectively is necessary. In addition, the way of searching effective wrapping processes which satisfy the demand of user is necessary to fulfill the requirement of beginners as assistant means of paper wrapping.

In this paper, we propose an approach to support the paper wrapping activity by managing the processes of wrapping a target object with a piece of papers. In order to support paper wrapping, it is necessary to represent, generate and compare various wrapping processes. Therefore, in our approach the wrapping processes are managed hierarchically with tree structure. In this way, we can also delete non-effective wrapping processes easily and it is easy for users to check out various wrapping conditions. Folding operations are constituted by simulating in virtual space, called an internal model, that describes both information of an object and deformation processes of wrapping paper, and then the tree structure is constructed by generating folding operations over and over. In addition, we propose a method for searching effective and feasible wrapping processes based on the demand of user. The purpose of this study is to develop a system that supports the paper wrapping visually by CG animation in wrapping processes.

2 Framework

2.1 Wrapping Process

Even when a target object and a piece of wrapping papers are given, the wrapping process is not decided uniquely. In fact, individuals wrap materials with the different ways in different applications even if they used the same object and the same wrapping paper. So, the required wrapping processes depend on the demands of users.

In order to support the paper wrapping activity, it is necessary to design an effective wrapping process which satisfies the following conditions: there are no exposed object faces and visible backsides of a wrapping paper, and the process is physically foldable. Thus, it is necessary to analyze whether designed wrapping processes are effective or not and well manage various wrapping processes. To deal with these issues, a powerful method for specifying the deformation processes of wrapping paper and managing the effective knowledge is necessary.

The conditions which are used to judge whether the given crease pattern is feasible and/or foldable in three-dimensional space were argued in Origami mathematical fields[1]. The conditions are called *fold-ability*, and are dependent on the position and folding angle of the crease lines on crease pattern. Although only the necessary condition for fold-ability has been provided, the sufficient condition has been not so. Thus,

there are a lot of crease patterns which cannot be folded although they satisfy these conditions.

We construct a structural state space which represents wrapping processes on a tree structure, called *stage tree* to manage the effectiveness of wrapping processes for users. Each node in the tree represents one stage in each wrapping process. Applying the folding operation, which corresponds to an edge of tree, to each stage, a new stage is generated. Wrapping processes are designed by constructing a stage tree while generating folding operations. In this way, many effective or non-effective wrapping processes are managed hierarchically and users can view the wrapping condition which each stage attached with. In addition, it is easy to cut out the stages which are not necessary for users. So, it is allowed to judge *fold-ability* by actually generating folding operations in a virtual space.

2.2 Definition of Wrap-Folding

A wrapping paper is folded toward an object face without conflictions of faces which compose a wrapping paper. Fig. 1 shows some operations that avoid conflictions along the object face. Especially, the folding operations in Fig. 1(b) and Fig. 1(c) are peculiar to the paper wrapping. However, because there are various folding operations used in the paper wrapping, it is difficult to define the operations. Therefore, we restrict the folding operations to ones in Fig. 1, and define these folding operations as *wrap-folding*. Wrap-folding is three patterns of folding operations in Fig. 1 according to both the condition of object and wrapping paper.

The procedure for generating wrapping operations judges conflictions of faces, generates crease lines and rotates the divided faces. The internal model is renewed after the generation of folding operations. The renewal of the internal model is based on the moved face by the generation of wrapping operations. In Section 3, the method for generating wrapping operations is described.

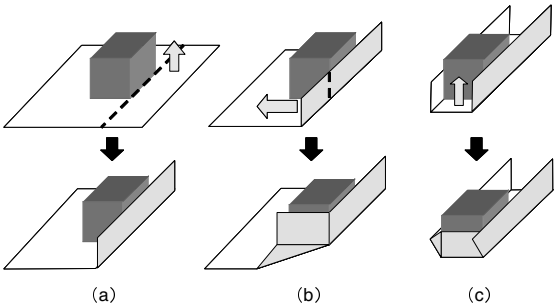


Fig. 1. An example of wrap-folding

2.3 Internal Model

State of Wrapping. In order to design wrapping processes by using the information about a target object and a wrapping paper, we need an internal model which can represent stepwise states of wrapping. We use the Origami model as internal model for

paper wrapping, which Miyazaki, et al. [4] have already proposed. This model consists of vertices V, edges E and faces F as the basic data elements. An edge $e_{i,j} \subseteq E$ has information about its coordinates and the angle $\theta_{i,j}$ between two faces f_i, f_j. If $\theta_{i,j} = \pm\pi$ then the edge(crease) means valley/mountain folding. In addition, the model has a data structure which can represent the overlapping faces on the same plane. An example of this structure is shown in Fig. 2. The structure groups the faces on the same plane and holds the order of overlapping by a face list ($f_2 \rightarrow f_3$). A wrapping operation is applied to faces on the same plane. Each face is stored at a face list in the order of the normal vector calculated from the object face on the same plane.

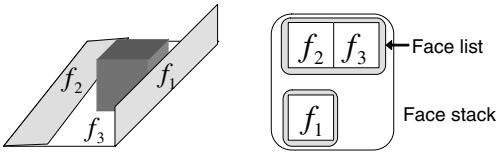


Fig. 2. A structure for describing overlapping faces

Stage Tree. When the information about an object and a wrapping paper is given, there are a lot of possible wrapping processes. Moreover, searching the wrapping process which satisfies the demand of user from many processes is required. In Fig. 3, a tree structure (stage tree) in which nodes represent wrapping stages and edges represent folding operations is shown. The tree structure obtains the relationships among wrapping stages and a record of wrapping processes. The route that arrives from the root to the leaf shows one sequence of wrapping process. A new stage is generated by applying any folding operation to an existing stage.

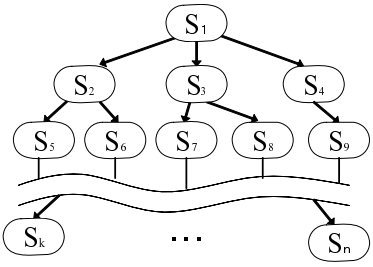


Fig. 3. Decision tree in wrapping processes

2.4 User-Interactive Pruning

The number of stages in a stage tree is very large, if the folding operations are applied to all foldable faces. From the viewpoint of paper wrapping support, since it is important to provide effective wrapping processes which satisfy a demand of user, it is appropriate to cut off unnecessary wrapping processes which a user does not require. However, we encounter difficulties when incomplete wrapping stages are evaluated, because we cannot

predict the final wrapping condition from the incomplete wrapping stage. Namely, predicting the final wrapping condition is equivalent to calculating the wrapping process completely. In addition, there are various demands from users for wrapping support: e.g. appearance, degree of difficulty, minimum required size of a wrapping paper, etc. There are no wrapping processes which satisfy various demands at once.

In our method, a user determines evaluation criterions for generating a stage tree. In other words, the wrapping processes not required are cut off from a stage tree. Some wrapping processes which satisfy a demand of user should be provided, even if the most effective wrapping process is not specified. Additionally, as a depth of stage tree is deeper, a stage on the way exerts less effect on final wrapping condition. Therefore, a user can determine the depth of *sub-stage tree* which the folding operations are applied to all foldable faces. All stages which belong to the sub-tree are never cut off. In this way, the accuracy of tree search is determined and the most effective wrapping process does not have to be searched exactly. In Section 3.4, our pruning algorithm is described in detail.

3 Method of Generating Stage Tree

3.1 Processing Flow

Applying a folding operation to a certain stage, all the feasible stages are generated. The process of generating stage tree is divided into three phases. The flowchart is shown in Fig. 4.

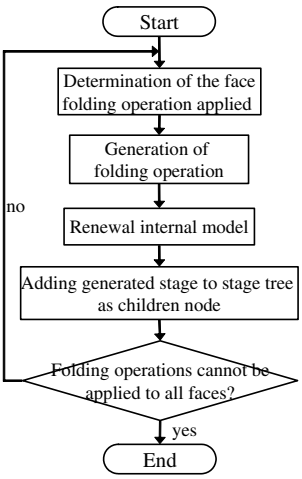


Fig. 4. Flowchart of generating the feasible stages from a certain stage in one wrap-folding operation

First, a target face for a folding operation is decided in each stage. All target faces should be foldable at least for one face of object. There is a plural number in the faces that can become a target face in each stage, and a stage tree branches off by depending

on how to choose these faces. Second, the folding operation is actually generated. The folding operation is constituted by generating crease lines, dividing faces and moving the divided faces. We show the method for generating crease lines in detail in Section 3.3. Finally, the face list in our internal model is updated. Whenever a folding operation is generated, a face list is updated.

As a wrapping stage becomes more complex, the number of the faces stored in a face list increases. In this case, each folding operation is applied to all stored faces in a face list in which target face is stored in the same way. These processings are repeated as far as there are the faces in which folding operations are calculated.

3.2 Extraction of Folded Faces

In this section, we show the method for determining face to which a folding operation can be applied. In this method, the face is judged based on inclusive relation among faces. Concretely speaking, a wrapping face f_p can be folded onto an object face f'_o only if the following judgment conditions are satisfied at the same time:

1. There are neither f_p nor f'_o on the same plane, and
2. There is an edge constituting f'_o which intersects with f_p.

f_p which satisfies these conditions for at least one f'_0 is added to a target face list that is a temporary list. In this research, on the basis of all the possible combinations of the

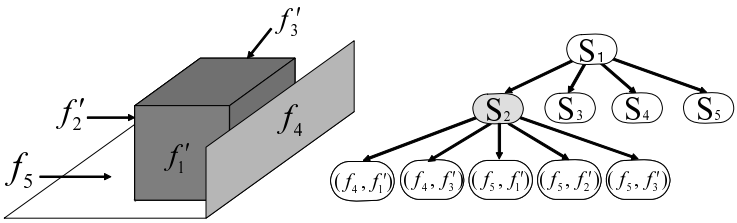


Fig. 5. An example of judging the wrapped faces

folded face f_p and wrapped face f'_0, each stage is generated by a folding operation. The following equation shows all the combinations in an example of Fig. 5.

$$(f_p, f'_o) = \{(f_4, f'_1), (f_4, f'_3)\}, \{(f_5, f'_1), (f_5, f'_2), (f_5, f'_3)\}$$

In this case, the number of combinations between f_p and f'_o is 5. In the stage tree of Fig. 5, the grayed node S_2 is calculated, and five child nodes are generated from S_2.

3.3 Method of Generating Folding Operation

The flowchart expressing the basic processing of generation method for wrap-folding is shown in Fig. 6.

In wrap-folding, if f_p can be folded without conflicting all other faces, a folding operation is generated using the rotational transformation of f_p around the axis

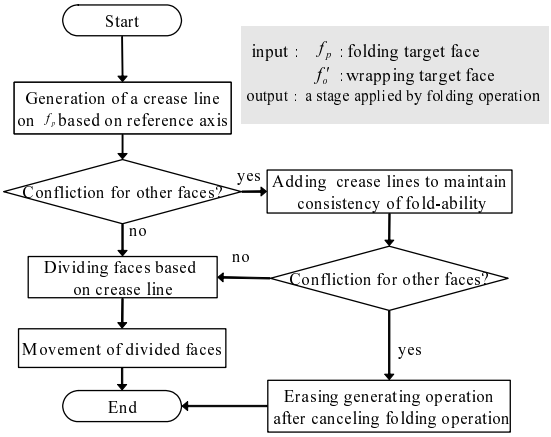


Fig. 6. Flowchart of generating wrap-folding operation

corresponding to the crease. Otherwise, by adding other folding operations, the collision among faces is evaded like Fig. 1(b) and Fig. 1(c). In this section, we show the processing in detail.

Definition in Generating Crease Lines. In Fig. 7, we define a folded face of wrapping paper as f_p, a wrapped face of object as f'_o, face that conflicts with f_p when f_p is folded as a confliction face, an edge of f'_o which contacts to f_p as e'_b and an edge that will be erased after the folding operation as a unification edge. Furthermore, an intersection between e'_b and a unification edge is defined as a confliction point. In this method, by using f_p and f'_o a confliction face is computed and crease lines are generated. There are one or two confliction points which depend on the relationship between f_p and f'_o.

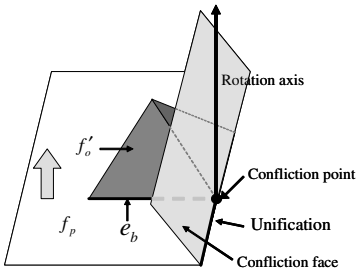


Fig. 7. Definition for generating one wrap-folding

Generating Crease Lines. In this section, we show a method of generating crease lines. In the case that N object edges $\{e'_i \mid 0 \leq i \leq N-1\}$ connect with confliction point, the processing of generating crease lines is as follows. e'_i is a linear list of the object

edges connected from e'_b sequentially around a confliction point. Furthermore, β_i is the plane angle between e'_i and e'_{i+1}, and satisfies:

$$\sum_{i=0}^{N-1} \beta_i < 2\pi \tag{1}$$

1. e'_b is defined as the base line and the crease line e_0 having the folding angle θ'_0 provided from e'_b is generated.
2. The following procedures are repeated for all $i's$ $(0 \leq i \leq N-1)$.
   (a) The crease line e_{i+1} is generated at a position of the angle γ_i from a base line calculated by the following equation:

$$\gamma_i = \sum_{j=0}^{i} \beta_j \tag{2}$$

   (b) $\theta_{i+1} \leftarrow \theta'_{i+1}$.
3. The crease line e_N having the folding angle π is generated at a position of the angle α from a base line provided by the following equation:

$$\alpha = \frac{2\pi - \gamma_{N-2}}{2} \tag{3}$$

4. A line between f_p and confliction face are divided at the confliction point, and then unification edge and crease line e_{N+1} are generated.

An example of generating crease lines is shown in Fig. 8. First, the base line is specified by the extended line of e'_b, and the crease line e_0 having the folding angle θ_0 is generated. Second, the crease line e_1 having the folding angle θ_1 is generated at a position of the angle β_0 from the base line. The angle θ_1 equals to the angle of e'_1. In the same way, the crease line e_2 having the folding angle $\pi - \theta'_2$ which depends on the folding angle θ'_2 of e'_2 is generated at a position of the angle β_0 from the base line. Finally, it is necessary to generate the crease line in order to maintain the consistency of folding ability. This crease line is generated at a position of the angle α from the base line by using Equation (3). The folding angle is π. Then, the folding angle of the unification edge is made to be 0 by dividing a line between f_p and confliction face. We show the state of wrapping before and after generating crease lines in Fig. 8. The unification edge is abbreviated in Fig.8(b).

3.4 Pruning Algorithm

Some wrapping processes are generated based on the demand of user. The appropriate numerical wrapping processes which satisfy the demand are provided to users. The hill-climbing method is used as the basic tree constructing method, and this processing is divided into two phases. First, sub-stage tree whose depth is lower than d which user determines is generated. Next, each generated leaf stage is viewed as an initial wrapping state, and the construction for these stages is performed using hill-crimbing method.

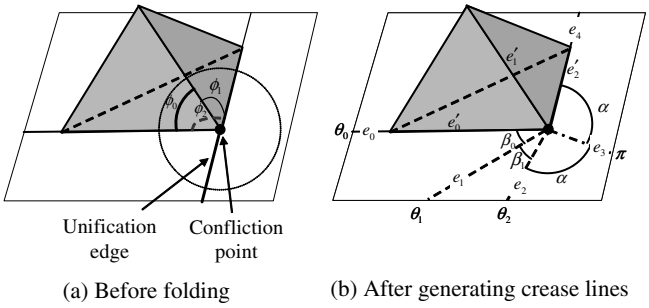


Fig. 8. Wrap-folding operation

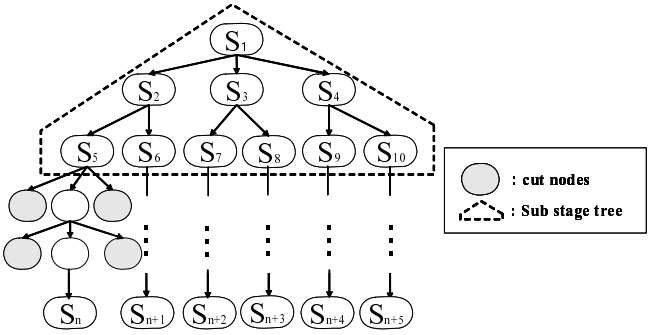


Fig. 9. Stage tree when the depth of sub-stage tree is 2

Consequently, the number of generated wrapping processes satisfies the demand of user. Fig.9 illustrates the distribution of search area if the depth of sub-stage tree is 2.

We use two variables E_{comp} and E_{back} to evaluate the difficulty of a sequence of folding operations and the appearance of complete wrapping states. Since the evaluation with these criteria is equivalent to the evaluation for edges of stage tree, it is possible for incomplete wrapping stages to be evaluated. The evaluated value of a stage E is

$$E = \frac{\alpha E_{comp}}{2} + \frac{\beta E_{back}}{N} \quad (\alpha + \beta = 1) \tag{4}$$

E_{comp} is the number of confliction points when wrap-folding is applied to the previous stage. E_{back} is the number of face-groups which have viewable backside in generated stage and N is the total number of face-groups. These values are added to the evaluation value of the parent stage, and the value is viewed as the evaluation value of the target stage. Because not only the final wrapping condition but also the wrapping condition on the way is important as for the paper wrapping, it is proper to search effective wrapping processes with these end-points. Each coefficient in Equation (4) is determined by a user in order to make the generated wrapping processes effective.

4 Experimental Result

Our approach for designing a sequence of wrapping processes was tested to validate the evidence of availability. We performed our experimentation from a viewpoint of construction of stage tree and the number of generated stages.

4.1 Construction of Stage Tree

From the state of both an object and wrapping paper in Fig. 10, a stage tree is constructed with our method. In order to calculate wrapping processes, all the feasible ways of folding operation are constituted. As a result, 16306 stages are generated and 31 wrapping processes are generated as the effective wrapping processes. Feasible and consistent folding operations which fold the folded face toward the wrapped face were constituted in 3-D virtual space. When there was a confliction face, the confliction face and the confliction point are computed from the relation between the object and wrapping paper, and then folding operations are applied without confliction of faces. One of the effective wrapping process is shown in Fig. 10.

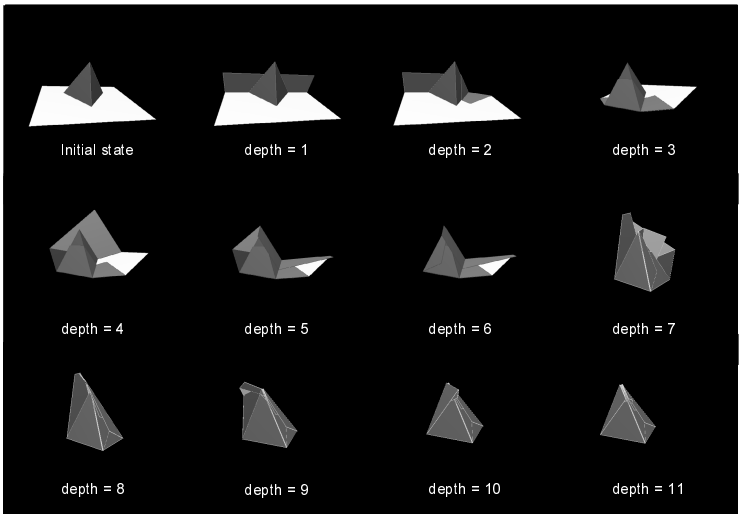


Fig. 10. One result of wrapping processes

4.2 The Number of Generated Stages

A stage tree is constructed with proposed pruning algorithm from the same initial state to test the number of all generated stages and effective wrapping processes. The values $\alpha = 0.5, \beta = 0.5$ are used as evaluated weight factors. In order to compare the number of generated stages, we estimate the difference of the sub-stage tree's depth d in which all search methods are applied. As a result, the number of all generated stages is reduced with proposed pruning algorithm. However, there are no effective wrapping processes

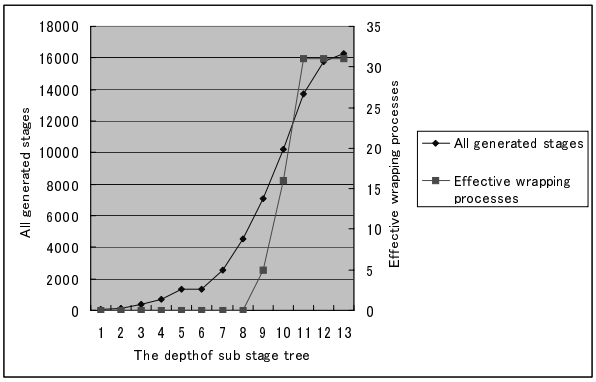


Fig. 11. Experimental results related to the number of generated stages

when the depth d is less than 9. The different number between all generated stages and effective wrapping processes are plotted in Fig. 11.

4.3 Discussion

We proved that many unnecessary stages are reduced with our proposed pruning algorithm through experiments. On the other hand, effective wrapping processes are also reduced. Effective wrapping processes were not generated in this experiment when the depth of sub stage tree is less than 9. Depending on the value that the user determined, there is the case that an effective wrapping process is not generated at all. So, we arrive at our conclusion that we need a more effective evaluation method which makes it possible to predict whether the wrapping process is complete.

In addition, the wrapping processes are regarded as different processes when the order of wrapped face is different, even if generated last wrapping state is equivalent. Therefore, the difference between generated wrapping processes was small, and many wrapping processes whose last wrapping condition was equivalent were generated differently in this experiment. From a viewpoint of paper wrapping support, most effective wrapping process does not have to be specified closely. All we need to do provide the guidance principle of wrapping processes which satisfy the demand of user. It seems that it is effective to use the information about the crease pattern on a wrapping paper which is mainly used in research field about Origami. Because the crease pattern has only the crease line information of the wrapping paper, and the information about overlapping faces is not represented. In other words, it means that the crease pattern has only the wrapping essential information. We can treat only the essential property of the wrapping processes which are necessary point for paper wrapping support by using the information about crease pattern as a basic evaluation criterion.

5 Conclusion

We have proposed an approach to represent and design wrapping processes in paper wrapping as support. As the main result, the method for managing multiple wrapping

processes with a tree structure and formalization of folding operations for paper wrapping were proposed. The proposed method made it possible to design various foldable wrapping processes and represent a sequence of wrapping processes by CG.

In addition, the pruning algorithm for a stage tree is explained. The evaluation criterions are determined by user and unnecessary wrapping processes are cut off. Many wrapping processes which were unnecessary for user were reduced, but some effectively generated wrapping processes were also reduced. The result of our experiment clearly shows that the method for predicting the final wrapping condition is necessary.

In this paper, we have considered the wrapping process by focusing on the faces of both the object and wrapping paper in generating folding operations. Although our approach makes it possible to design multiple wrapping processes, generated stage tree has many useless information and we does not consider the initial position of wrapping paper and object. From the viewpoint of paper wrapping support, we need the method for analyzing guidance principle of generating stage tree so to decrease useless information and make generated wrapping processes effective. A foreseeable extension of this research would be to design wrapping processes using information of crease patterns. To use information about crease pattern is also useful to estimate optimally initial position of wrapping paper and object.

References

1. Belcastro, S., Hull, T.: Modeling the Folding of Paper into Three Dimensions Using Affine Transformations. Linear Algebra and its Applications, 466–476 (1999)
2. Hull, T.: On the Mathematics of Flat Origamis. Congressus Numerantium (100), 215–224 (1994)
3. Kawasaki, T.: R( )=1. In: Proc. of the 2nd International Meeting of Origami Science and Scientific Origami, vol. 3, pp. 31–40 (1994)
4. Miyazaki, S., Yasuda, T., Yokoi, S., Toriwaki, J.: An ORIGAMI Playing Simulator in the Virtual Space. The Journal of Visualization and Computer Animation 7, 25–42 (1996)
5. Shimanuki, H., Kato, J., Watanabe, T.: Constituting Feasible Folding Operations Using Incomplete Crease Information. In: Proc. of IAPR Workshop on Machine Vision Applications, pp. 68–71 (2002)

Hangeul Learning System

Jae won Jung and Jong weon Lee

Mixed Reality and Interaction Lab, Sejong University,
98 Kunjadong, Kwangjinku,
Seoul 143747, Korea
deeploveme@hotmail.com, jwlee@sejong.ac,kr

Abstract. This system is not only for a foreigner but also for everyone in Korea who doesn't know Hangeul (Korean language). It is difficult to study Hangeul themselves without any helper. This paper presents the AR based system that helps people to learn basic Hangeul letters and pronunciations in their home without any helper by applying the characteristics of consonant and vowel. We also suggest the Word Studying Methods using the proposed system. At this time, it is developed based on the pattern matching function of ARToolKit, we plan to improve the system by applying the character recognition function.

Keywords: Augmented Reality, Edutainment, Korean Language Learning System.

1 Introduction

The language learning environment has been sustained with text–oriented environment for many years. It has been changed slowly to match the today's digital era. The recent technology trend is not simple information transmission but overall information transmission (e.g. ask for the sight, the sound, mutual communication). That is the latest trend of acquisition of knowledge in the education field. Although some traditional instructions use multimedia, but it has limited users' participation. It did nothing else than keyboard typing or mouse click. In this paper, to overcome these limits and to improve the current language learning systems, we propose a new approach to assist learning the Hangeul language using Augmented Reality.

Korean language 'Hangeul' is one of the most interesting languages. Hangeul creates words by combining the basic buildingblocks, Jamo. At least two and often three of Jamo are placed on the pattern to form a word. Depending on the pattern and Jamo, users can create various words. Since this type of arrangement is not usually used in other language, foreigners may have difficulty to learn Hangeul. We propose the Hangeul Learning System using AR to overcome this problem

This is followed by related works in section 2 and details of the main components in Sections 3, and results illustrating for several different scenarios are presented in Section 4. Section 5 is Conclusion.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 126–134, 2008.

2 Related Works

The most well known system for learning languages using AR is the system that can learn Japanese Kanji [5]. After this system was introduced, few systems that could assist users to learn languages have been developed. There are also few education systems using AR. There are systems to teach molecular structure in the field of chemistry or atmosphere circulation in earth science. They are designed to use in school and use visual information to increase interests of users. In this section, we present the Japanese Kanji learning system and two science education systems.

2.1 Kanji Education System

It was made for helping users to learn Japanese Kanji [5]. Users learn Kanji characters by playing the matching games developed using handheld Augmented Reality technologies. A user supposes to find the corresponding answer when the question is given on the screen.

2.2 Chemistry Education System

Many people may think ARbased education systems are only for children. But the presented system in this section is not for children but for welleducated people. Users can see the 3D structure of the molecular using the system [1].

The system is formed with browser, tagtoggle, cleaner and molecular marker. Users can see molecular structure and atom they want. Traditional text based chemistry education can not see as 3D structure of a molecular, but the system shows 3D structure of the molecular and helps users to understand the molecular structure better.

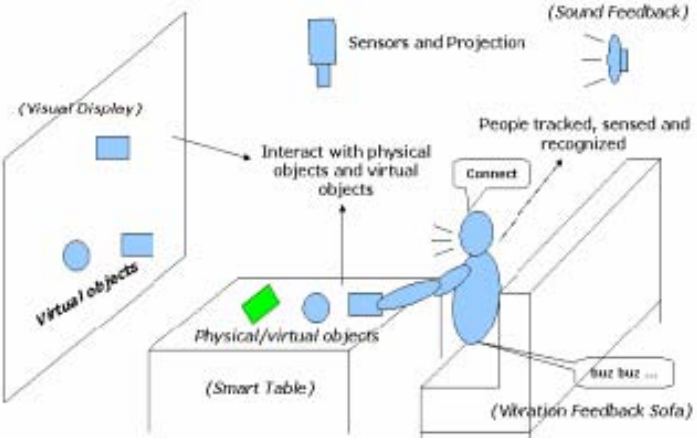


Fig. 1. Expected Hangeul Learning System [4]

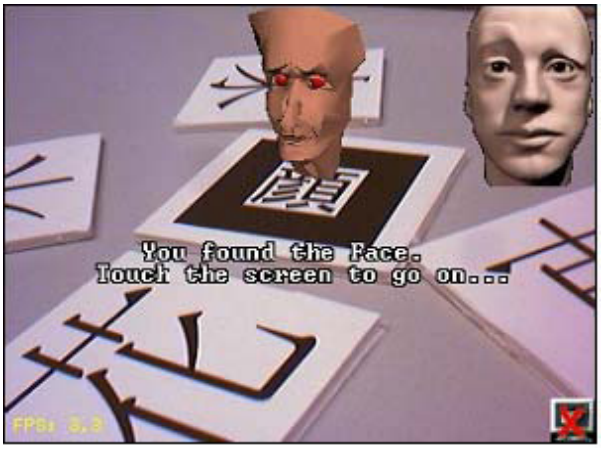


Fig. 2. Kanji education system

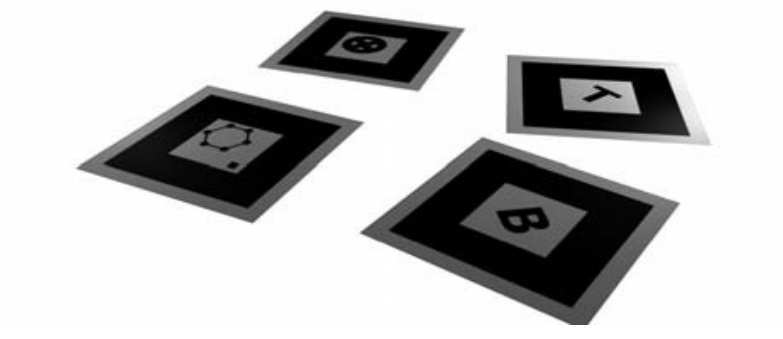
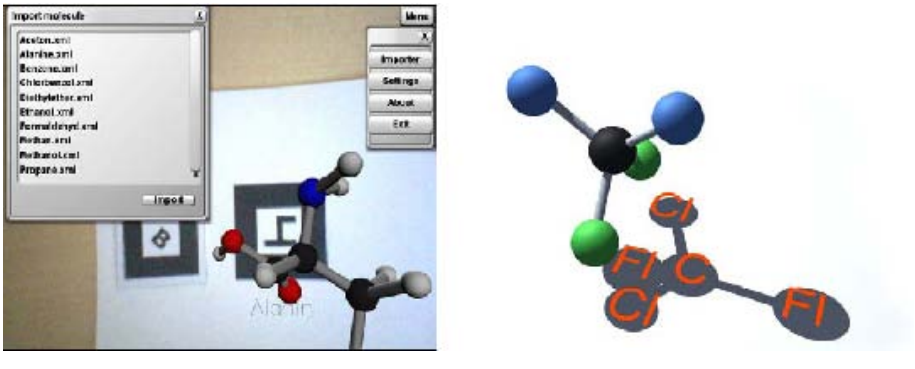


Fig. 3. The function cards



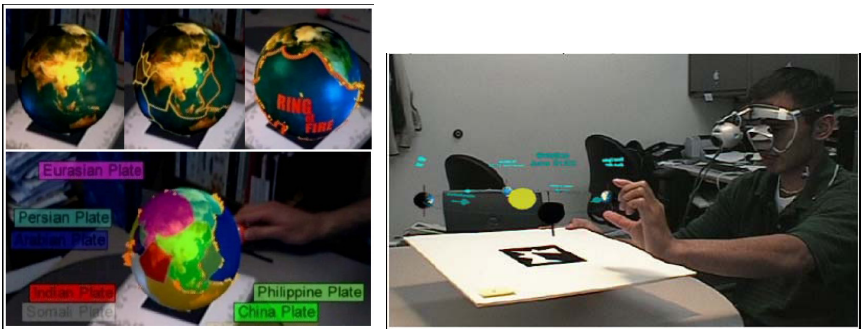
(a) (b)

Fig. 4. The chemistry education system (a) GUI overview with menu button and main menu and molecule structure (b) 3D visualization of the molecular structure

2.3 Weather Education System

This system shows principles that form rain and cloud as shown in Figure 5 [4]. The system is good for children. It can simulate the water circulation through scenarios. The system contains auditory part and visual part. The system provides a rain sound to improve the learning experience. The system also induces user's participation in sc enarios. Users can modify the position of a marker to interact with the system.

Fig. 5. Weather education system



(a) (b)

Fig. 6. (a) Geological features; (b) Using Augmented Reality for Teaching EarthSun Relation ships to undergraduate Geography Students

Another similar system is the Solar-System and Orbit Learning system [2]. Using this system, users can understand why earthquake occurs and learn volcano topography and name of nature of soil. Users can also learn the relationships between Earth and Sun by using the system shown in Figure 6 (b) [3]. The relationships are represented in 3D, so users can view at various positions and orientations.

3 The Proposed System

The proposed system can be easily installed to a PC with a web camera. The system can be easily used by anyone at home and at school. This paper proposes the effective language learning system that utilizes visual and auditory senses of users. The proposed system could help people to learn the structure of Hangeul. We introduced an interesting approach in treating consonants and vowels of Hangeul, Jamo. Each character is recognized by combining given constants and vowels, and its corresponding 3D objects and sounds is presented to a user.

3.1 Approach

The Hangeul Learning System is able to recognize and to express variety of characters using the characteristics of Hangeul. When a person uses this system, the marker representing each character must be registered. For example, the word '가' can be recognized by the system when the corresponding marker is registered. However, the user put two markers, 'ㄱ' marker and 'ㅏ' marker together, to create the word as shown in Figure 7.

The same two markers can be combined to create different words such as '고', '나' and '거'. Since 'ㄱ' marker could rotate, it would be used for words requiring the marker 'ㄴ'. Likewise, 'ㅏ' marker could be used for words requiring 'ㅓ' or 'ㅗ' markers. This approach can reduce the number of markers needed in the system. Learners can learn many words using few markers. Figure 8 and Figure 9 describe the approach by using characters of English and Japanese to help readers understand the concept.

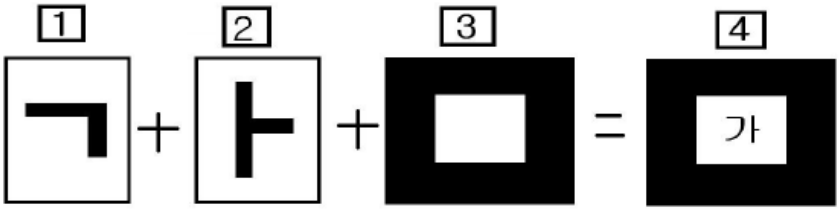


Fig. 7. The way to combine markers to form a word

3.2 Functionalities

The Hangeul Learning System consists of five areas as shown in Figure 10. When a user put markers on the area 1, the system recognizes the word and shows the word by displaying the corresponding text, image or 3D object at the area 2. The system recognizes words by registering corresponding markers in the system and by using the pattern matching part of ARToolKit . When the user touches the marker on the area 4, the corresponding sound is produced. When the user touches the marker on the area 5, the phonetic sign is displayed on the area 3.

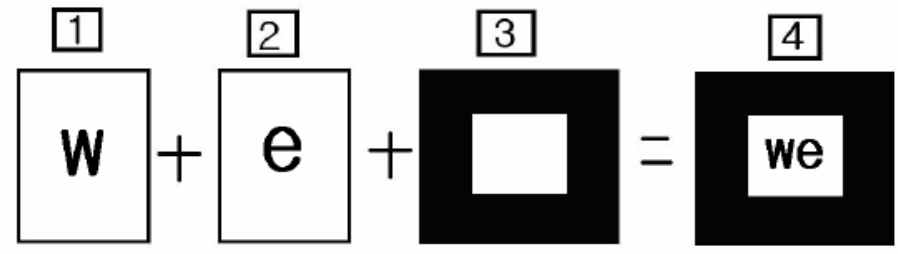


Fig. 8. The way to combine markers to form a word (English)

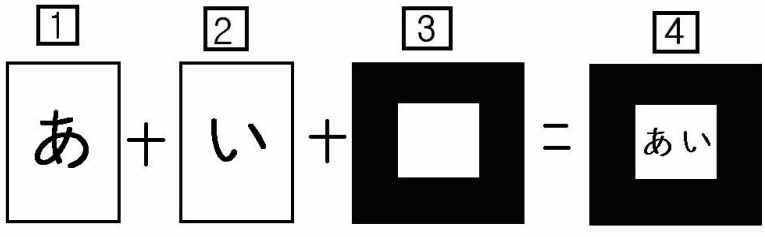


Fig. 9. The way to combine markers to form a word (Japanese)

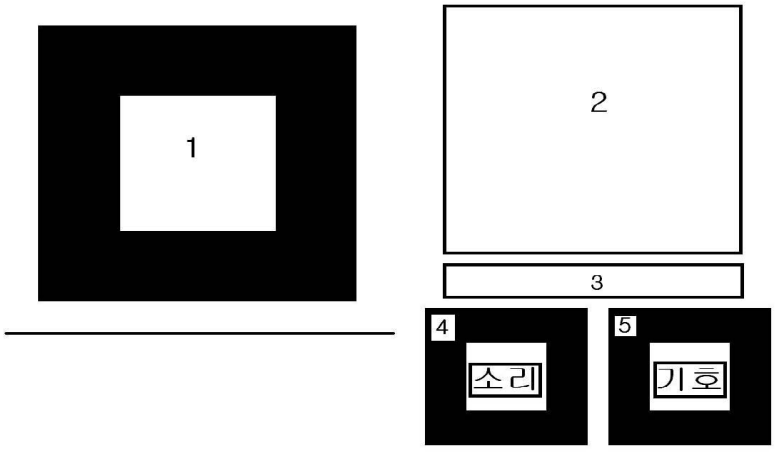


Fig. 10. Five areas used in the Hangeul Learning System

4 Experiments

Figure 11 (a) is initialization screen. A user put ‘ㄴ’ and ‘ㅏ’ to create ‘나’ (Figure 11 (b)). The marker ‘ㄴ’ can be also used for ‘ㄱ’ by rotating it as shown in

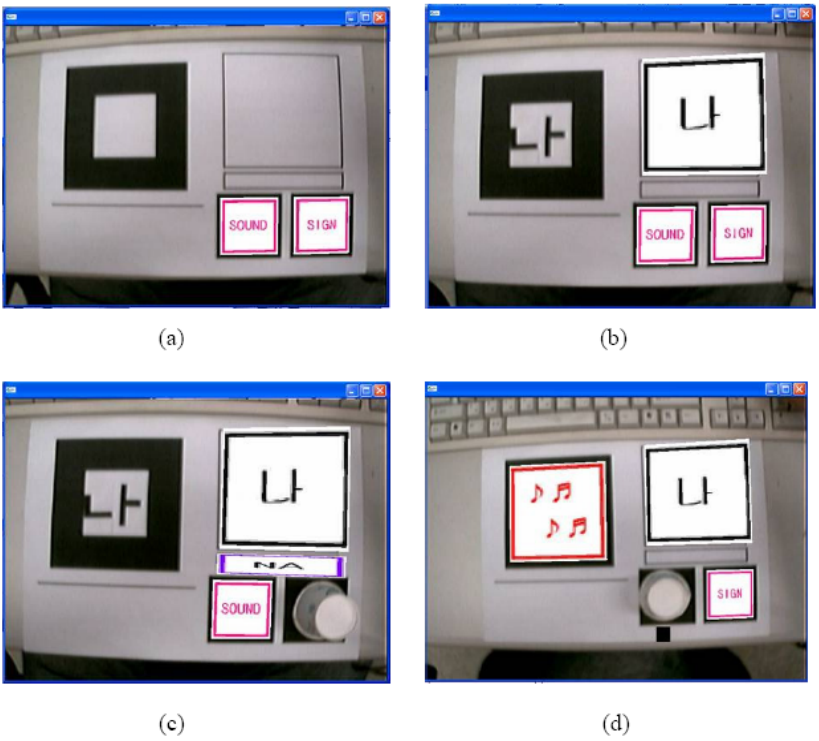


Fig. 11. '나' is created by markers rotated 'ㄱ' and 'ㅏ'

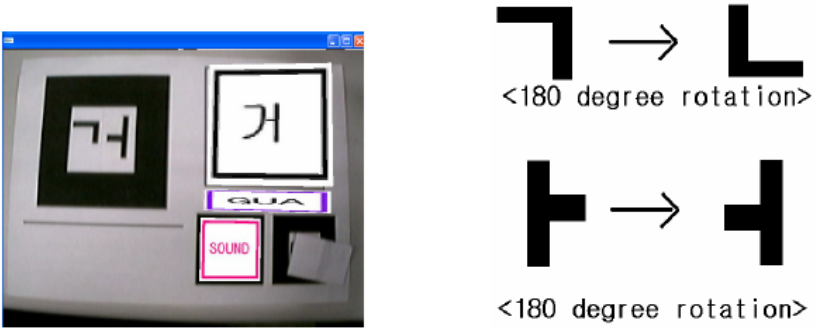


Fig. 12. '거' is created by markers 'ㄱ' and rotated 'ㅏ

Figure 12. Figure 11 (c) and (d) illustrate the outcome of touching each operation area, the area 4 and the area 5 in Figure 10.

The example in Figure 13 illustrates how related image and 3D model are used in the system. The usage of images and 3D models could help users to nderstand the word easily.

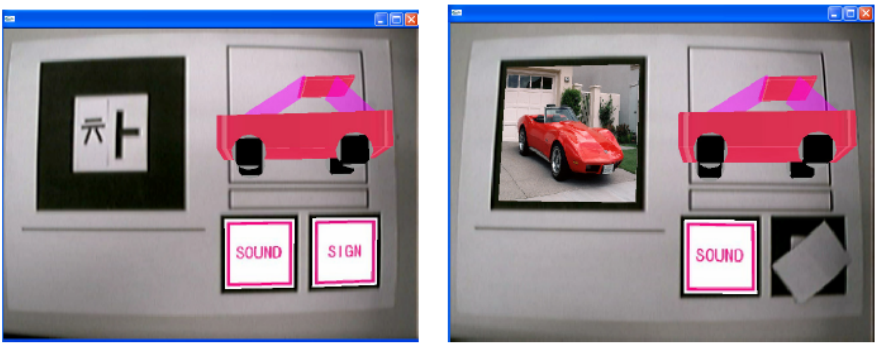


Fig. 13. An image and a 3D model used to represent the given word

5 Conclusion

In this paper we described the Hangeul Learning System that was developed based on Augmented Reality. The system has two main characteristics. The first one is creating the words by using fewer number of markers than the number of constants and vowels of Hangeul. The second is providing visual and auditory outputs to users. For visual outputs, 2D images and 3D models are used for the current system. We will plan to add the functionality that can provide related animation to users. As a result Hangeul Learning System can offer learners better ways to learn basic characters of Hangeul. The current system also has limitations. The current system only deals with words with one character. We will plan to provide users to learn words with more than one character. To reduce the number of markers registered, we plan to use the character recognition approach instead of the pattern matching approach. Since we create words with the given marker, the character recognition approach could be reliable.

Acknowledgments

This work was sponsored and funded by Korea Game Development & Promotion Institute as Korean government project (Ministry of Culture and Tourism).

References

[1] Almgren, J., Carlsson, R., Erkkonen, H., Fredriksson, J., Møller, S., Rydgård, H., Österberg, M., Fjeld, M.: Tangible User Interface for Chemistry Education: Portability, Database, and Visualization. In: SIGRAD 2005, pp. 19–24. Linköping University Electronic Press (2005)

[2] Woods, E., Billinghurst, M., Looser, J., Aldridge, G., Brown, D., Garrie, B., Nelles, C.: Augmenting the science centre and Museum Experience. In: Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia (GRAPHITE 2004), pp. 230–236. ACM Press, Singapore (2004)

[3] Shelton, B.E., Hedley, N.R.: Using Augmented Reality for Teaching Earth-Sun Relationships to Undergraduate Geography Students. In: Augmented Reality Toolkit, The First IEEE International Workshop, pp. 8–16. IEEE Press, Germany (2002)
[4] Kim, J.H.: Develop experience learning contents and research the application of a field based Augmented Reality. In: Korea education & research information service (keris), seoul (2005)
[5] Wagner, D., Barakonyi, I.: Augmented Reality Kanji Learning. In: Proceedings. The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, pp. 335–336. IEEE Press, Los Alamitos (2003)

An Ajax-Based Terminology System for E-Learning 2.0

Xinchun Cui*, Haiqing Wang, and Zaihui Cao

College of Info. Tech. & Comm., Qufu Normal University, Rizhao 276826, P.R. China
{cxcsd,whqet,czhhn}@126.com

Abstract. This paper presents the design, implementation and evaluation of a terminology system, which is an e-learning 2.0 community aiming to provide users timely and resultful help when they encounter a new term in their e-learning. It consists of term display module, term add module, term modify module and term query module. By introducing Ajax technology, it outgoes other terminology systems by better interaction, higher efficiency and prompt revisability.

Keywords: Terminology System, E-learning 2.0, Interaction, Ajax.

1 Introduction

A terminology system is to provide timely help for a learner when he encounters unfamiliar professional new. This is becoming more and more indispensable with the development of the nonlinear organized web-based learning environment. There are several terminology systems available such as Chinese Computer Terminology System [1] and the basketball dictionary in website hoopchina[2], they have done a lot of fruitful work, yet e-learning based on the system may be hindered by the inherent disadvantages as follows.

1. Poor interactivity. Based on the traditional network design model [3], if a user queries a professional term, he has to stare at a blank screen for several seconds in a poor web condition because these systems must refresh the full page, which tremendously decreases his learning interest and interrupts his learning process.
2. Low efficiency. Most of current terminology systems are independent ones, that is, they are not committed to any web-based course. So, when a user encounters a new term in an article, he has to open a new page of a terminology interpreting system for help. That is time-consuming; furthermore, he may be misled by several such operations.
3. Low revisability of term database. The model of modifying term database of most current terminology systems is manual. So that it is a hard work to update the term database. Beside time-consuming, it is also error-importing. This leads to low revisability.

E-learning 2.0 (as coined by Stephen Downes[4]) takes a 'small pieces, loosely joined' approach that combines the use of discrete but complementary tools and web

* Corresponding author.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 135–146, 2008.

services - such as blogs, wikis, and other social software - to support the creation of ad-hoc learning communities[5]. This paper makes use of Ajax technology to create an E-learning 2.0 community. By introducing Ajax technology, the proposed terminology system outgoes others by better interactivity, higher efficiency and timely modification. Specifically, the following approaches are introduced to improve the system, 1) automatically add explanation to term, 2) search suggest, 3) save automatically, 4) preload page.

The rest of this paper is structured as follows. We initially present basic of terminology system, including shaping education by e-learning 2.0 and basic of Ajax. Next, the architecture of terminology interpreting system is presented. Afterwards, this paper provides the implement of the system. Finally, some concluding remarks and our vision for the next steps are presented.

2 Basic of the Terminology System

2.1 Shaping Education by E-learning 2.0

There are some very interesting changes going on in the world of e-learning. The changes in e-learning are being driven by two primary forces; the first force is a steady increase in the pace of business and information creation, the other is the advent of Web 2.0.

There are three Generations of E-Learning, E-Learning 1.0, E-Learning 1.3, and E-Learning 2.0.(see Table 1) E-Learning 2.0 is based on tools that combine ease of content creation, web delivery, and integrated collaboration. Creation of content can occur by anyone as part of their day-to-day life [6]. In essence, the expectation of E-Learning 2.0 is that sharing and learning becomes an organic action that is directed and driven by the learners. Learners are starting to explore the potential of blogs, media-sharing services and other social software - which, although not designed specifically for e-learning, can be used to empower students and create exciting new learning opportunities [7].

The proposed terminology system can provide learners good experience. Learners not only can query and see the definition of the term but also can modify or add the definition when they need it. Learners become their learn process's drivers and content creators. So, the system is an e-learning 2.0 app.

Table 1. Three Generations of E-Learning

	E-Learning 1.0	**E-Learning 1.3**	**E-Learning 2.0**
Ownership	Top-down, one-way	Top-down, collaborative	Bottom-up, learner-driven, peer learning
Access Time	Prior to work	In between work	During work
Delivery	At one time	In many pieces	When you need it
Content Access	LMS	Email, Intranet	Search, RSS feed
Driver	ID	Learner	Worker

2.2 Basics of Ajax

Asynchronous JavaScript and XML (Ajax), Ajax isn't a technology [8]. It's really several technologies, each flourishing in its own right, coming together in powerful new ways. Ajax incorporates:

1. Standards-based presentation using XHTML and CSS;
2. Dynamic display and interaction using the Document Object Model;
3. Data interchange and manipulation using XML and XSLT;
4. Asynchronous data retrieval using XMLHttpRequest;
5. JavaScript binding everything together.

Google Suggest and Google Maps are two examples of using Ajax to close the gap between desktop and web application. They take a great leap forth towards the richness of standard desktop applications. No longer are you forced to wait five seconds for the page to reload every time you click on something. Ajax applications change in real time. They can let you drag boxes around, they can refresh themselves with new information, and they can completely re-arrange the page without clearing it. And there's no special plug-in required. Ajax is just a style of design, one that milks all the features of modern browsers to produce something that feels fewer webs and more desktop.

The core idea behind Ajax is to make the communication with the server asynchronous, so that data is transferred and processed in the background. As a result the user can continue working on the other parts of the page without interruption. In an Ajax-enabled application only the relevant page elements are updated, only when this is necessary [9] [10].

Some of the characteristics of Ajax applications include:

1. Continuous Feel: Traditional web applications force you to submit a form, wait a few seconds, watch the page redraw, and then add some more info. Forgot to enter the area code in a phone number? Start all over again. Sometimes, you feel like you're in the middle of a traffic jam: go 20 meters, stop a minute, go 20 meters, stop a minute ... How many users couldn't endure too many error message and gave up the battle? Ajax offers a smooth ride all the way. There's no page reloads here - you're just doing stuff and the browser is responding.
2. Real-Time Updates: As part of the continuous feel, Ajax applications can update the page in real-time. Currently, news services on the web redraw the entire page at intervals, e.g. once every 15 minutes. In contrast, it's feasible for a browser running an Ajax application to poll the server every few seconds, so it's capable of updating any information directly on the parts of the page that need changing. The rest of the page is unaffected.
3. Graphical Interaction: Flashy backdrops are abundant on the web, but the basic mode of interaction has nevertheless mimicked the 1970s-style form-based data entry systems. Ajax represents a transition into the world of GUI controls visible on present-day desktops. Thus, you will encounter animations such as fading text to tell you something's just been saved, you will be able to drag items around, you will see some static text suddenly turn into an edit field as you hover over it.
4. Language Neutrality: Ajax strives to be equally usable with all the popular languages rather than be tied to one language. Past GUI attempts such as VB, Tk,

and Swing tended to be married to one specific programming language. Ajax has learned from the past and rejects this notion. To help facilitate this, XML is often used as a declarative interface language.

5. User first: Ajax was born with the idea to make user much more comfortable when he using web application by providing user with close-to-instantaneous performance, rich interfaces and tremendously improved user experience. This is of great importance to what just mentioned above.

Based on the characteristics of Ajax applications, we chose Ajax to improve user experience. In web-based learning environment, students' learning is highly dependent on learning environment, and vulnerable to outside interference. The traditional web application often leads students give up their learning because of its shortcoming for support student's learning, such as poor interactivity, unresponsiveness, simplistic interfaces, low usability. Fortunately, Ajax can overcome this.

3 The Architecture of Terminology System

The architecture of Terminology interpreting system (see Fig. 1) is based on a client-server platform model. The current form of Terminology interpreting system constitutes an open and flexible architecture with simple structure, which allows and supports the basic functionality that the platform is intended to offer. For this reason, the functionality of the system can be easily enriched by added another module. In addition, the Terminology interpreting system is characterized from openness due to the fact that is based on open technologies and international standards. More specifically, the implementation of the system is mainly based on (a) Ajax, including HTML, CSS, JavaScript, Dom, XML, etc., for generating a good interactive and friendly user client interface; (b) Java&JDBC, for implementing the server function of the system and realizing the communication between the server of system and database; (c)Mysql, for management the system's data.

In comparison with the previous Terminology System, this system uses Ajax to increase user experience in web-based learning environment. Exploiting with Ajax, the system takes the following measures to provide student with a student-centered web-based learning environment:

- **Automatically add explanation to term:** Automatically add links to terms in the page. As long as the term is in the database, the terminology system will add link to it automatically. When Tom is reading an article about java servlet in the network course, although he learned the term java before, the mean of it is always not clear. Learning may be interrupted if Tom doesn't know the mean of it. Fortunately, in this system, Tom can move mouse to the link to get the explanation of the word. The system will jump out a suggested frame, which includes the explanation of the term, the link of modifying the explanation of the term, the catalogue of the term and the correlative terms. (see Fig. 2.)

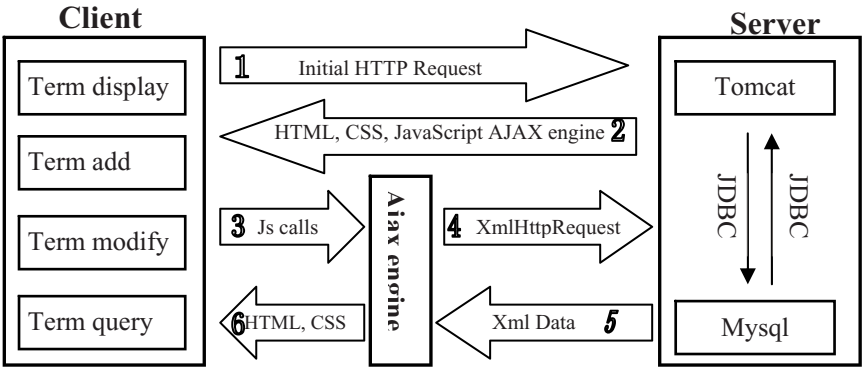


Fig. 1. Architecture of Terminology System

Fig. 2. Article page of network course

- **Search suggest:** The same as Google suggest, the interface is simply prominently features a text box to enter search terms. Everything appears to be the same until you start typing in the textbox. As you type, search suggest requests suggestions from the server, show you a drop-down list of search terms that you may be interested in. Each suggestion is displayed with a number of results available for the given term to help you decide.(see Fig. 3)The function may be very helpful for students when they type the “term” in the text box to search the mean of term, for he maybe don’t know the full spell of the term.
- **Save automatically:** Supposed that submission is failed after you changed the explanation of the term in term modify module. Don’t be depressed! The system has saved it automatically for you， you can find it in the drafts.

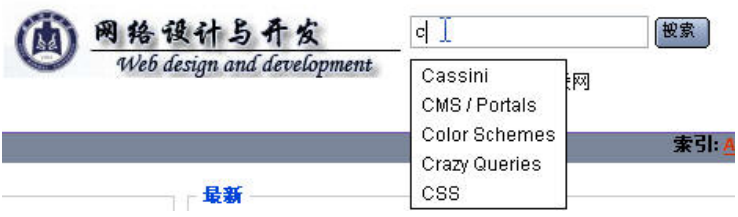


Fig. 3. Search suggest

- **Preload page:** You click the link to the next part of the article after you have read the first part of the article, you'll have to stare a blank page for five minutes sometimes even much longer in the traditional web system, and you'll lose your interest at the same time. The terminology system takes into account this point for you. When you enjoy the first part, the system guesses you'll read the next part and load it for you in background. So, when you click the link, the second part of the article is OK for you!

3.1 Client Side

As depicted in Fig. 1, the client side includes four main modules, term display module, term add module, term modify module, term query module. Detailed functional descriptions of the four modules are as follows:

- **Term display module:** There are two ways to display the interpretation of the terms. Firstly, the system automatically adds explanation to term, every term in users' reading articles will be added links to show explanations as long as the term is in database, as mentioned above. In addition, the style of the link is different with the others links in the system and decorated with dotted underline to help student to identify the link, which can be seen in Fig. 2. So, the student can better enjoy the article and gain knowledge in the reading. Secondly, learners can visit main page of terminology system (see Fig. 4), there are the list of new terms, the catalogs of terms and top 10 of contributors. Learners can see the explanations of terms, modify them, add them and query them.
- **Term query module:** If a learner finds a new term in reading and wants to know the mean of it, he may click a link added automatically by system to see the explanation; this is the result of query module. In another case, a learner is reading an article about XSL, he can type "XSL" in the below of user interface to query the explanation, he also can go to main page of system to get the answer. The system gives "search suggest" to help user to query the term.
- **Term modify module:** If learner find some areas of improvement for explanation of term when he/she see explanation, he can click the modifying link nearby to modify the explanation. Save automatically function can help learner save the page edited by learner. In order to guarantee the accuracy of explanation of term, student user's modifications need to be audited by teacher users. The explanations of terms in system are all audited.

Fig. 4. Main page of terminology system

- **Term add module:** Learner also can add term in system if find a term need to be made clear. Save automatically function also is used to prevent mishaps. Student user's modifications also need to be audited by teacher users.

3.2 Communication between Client and Server

The communication mechanism between client and server is asynchronous. As can be seen in Fig. 1, the processing flow is as follows:

1. Initial request by the browser – the user requests the particular URL.
2. The complete page is rendered by the server (along with the JavaScript Ajax engine) and sent to the client (HTML, CSS, and JavaScript Ajax engine).
3. All subsequent requests to the server are initiated as function calls to the JavaScript engine.
4. The JavaScript engine then makes an XmlHttpRequest to the server.
5. The server processes the request and sends a response in XML format to the client (XML document). It contains the data only of the page elements that need to be changed. In most cases this data comprises just a fraction of the total page markup.
6. The Ajax engine processes the server response, updates the relevant page content or performs another operation with the new data received from the server. (HTML + CSS)

3.3 Server Side

At server side, we choose Java and Mysql as functional implementing tools, because they are efficient and open. When the Sever receives a request form a client, it processes the task by searching the database. When the work is done, it transmits the answer to the client or tells it a failing message.

4 The Implement of the System

The core of implement of system is using Ajax to get rich user experience. More specifically, the core is achieve the function mentioned above, automatically add explanation to term, search suggest, save automatically, preload page etc. The general process to achieve the functions is the same. In this paper, take 'search suggest' as an example for illustrating the implement of system.

4.1 The Implement of Search Suggest

4.1.1 Data flow of Search Suggest

The data flow of search suggest module is shown in Fig. 5.[11][12] Firstly, A user input a key word in text box to search; secondly, send an XMLHttpRequest to server; thirdly, server transmits the request and returns data to client; at last, client transmits the return data.

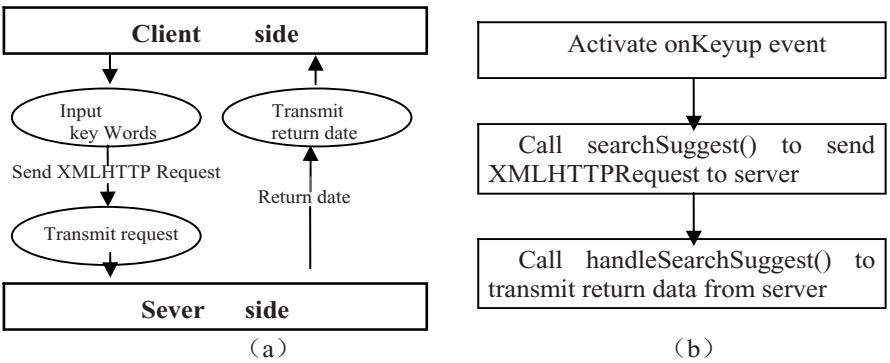


Fig. 5. (a) Data flow of search suggest module. (b) Flowchart of transmitting data in client.

4.1.2 Data flow of Client

Based on the basic principle of search suggest, the flowchart of transmitting data in client is depicted in Fig. 5.

4.1.2.1 Create XMLHTTPRequest. We use searchSuggest() to send a request to server. At first, define an XMLHTTPRequest object called searchReq, use crearteAJAXObj() to create searchReq. In the function createAJAXObj(), use "httprequest=new ActiveXObject("Msxml2.XMLHTTP");" to create the httprequest if the browser is IE, or use "httprequest=new XMLHttpRequest()" if the browser is the others. In the searchSuggest(), use searchReq.open() to send a request.

4.1.2.2 Transmit Return Data. You'll find it that handleSearchSuggest() is used to transmit return data. If the data is received successfully, use document.get ElementById('search_suggest') to get the div which id is search_suggest, and then use innerHTML to create the div which includes the return data. At the same time, use suggestOut() and suggestOver() to change the display to response the state of mouse

and use the function setSearch() to put the item choosed to the search text box. The code of handleSearchSuggest is as follows.

```
function handleSearchSuggest() {
  if (searchReq.readyState == 4) {
    var ss = document.getElementById('search_suggest')
    ss.innerHTML = '';
    var str = searchReq.responseText.split("\n");
    for(i=0; i < str.length - 1; i++) {
      var suggest = '<div
  onmouseover="javascript:suggestOver(this);" ';
      suggest += 'onmouseout="javascript:suggestOut(this);" ';
      suggest +=
  'onclick="javascript:setSearch(this.innerHTML);" ';
      suggest += 'class="suggest_link">' + str[i] + '</div>';
      ss.innerHTML += suggest;}}}
```

4.1.3 Flowchart of Server Transactions

In this system, Java and Mysql technology are used to process sever transactions. When sever receives a transaction, then the term database is connected and the transaction is submitted. A SQL statement such as “select title from suggest where title like '"+search+"%' order by title” is executed. The query result is usually put into a vector named as vDatan and is return to server. Flowchart of data transmitting in server part is shown in Fig. 6. The full code is in the appendix.

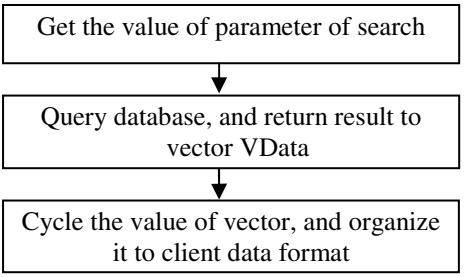


Fig. 6. Flowchart of server transactions

5 Application Note and Evaluation

5.1 Application Note

The terminology system is a part of the network course “Web Design and Development”, in the hope to help student conveniently find the relevant information

when he encounters a new term. In order to reach this, the system was designed simply and conveniently to use.

Take an example, when Tom is reading an article about java servlet in the network course, although he learned the term java before, the mean of it is always not clear. Learning process may be interrupted if Tom doesn't know the mean of it. Fortunately, in this system, Tom can move mouse to the link; the system will jump out a suggested box (Fig. 2). Tom can get the simple mean of "java" from the box or click the "detail" link to get more, he also can click the link "modify" to modify the mean of term if he thinks it can be improved. The related terms of "java", such as jsp, Ajax, javascript, are also displayed in the box; Tom can click the corresponding link to get the mean of the term which he is interested in.

During his reading, he may want to know the mean of asp, which is occurred to him. He can use the search input box in the right side of the bottom of page to search it (Fig. 2). During his inputting, he can get the help of "search suggest" (Fig. 3) and get a drop-down list of search terms that he may be interested in.

He also can go to main page (Fig. 4) of the system to search the term or modify the term. He also can add term in system if he finds a term needed to be made clear.

5.2 Evaluation of the System

The collection of the users' feedback and their elaboration will lead to an integrated, from both pedagogical as well as technological aspect, terminology system. The evaluation conducted was mainly focused on the usability and acceptability of the system. The aim was to evaluate the current functionality as well as the interface usability, in order to obtain results for future enhancements on the system.

The main questions asked were the following: (1) Is the system easy to use? (2) Is the system efficient to use? (3) Is the system subjectively pleasing? (4) Are you frequently modify or add the term? (5) Are you frequently interact with others? In order to survey the results, we set the answer is A. Strongly fit; B. mostly fit; C. fit; D. not fit.

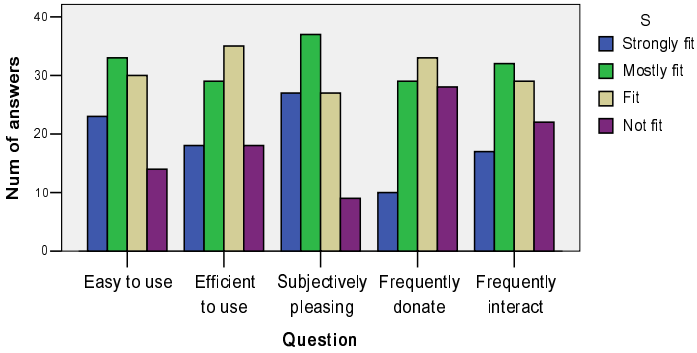


Fig. 7. The results of the evaluation

The Fig. 7 illustrates the result of the evaluation. We got 100 answers. For the first question, 86% of the users stated that the terminology system is easy to use, others thought the system is not easy to use. Concerning the efficiency of the system, 82% of the users (18% strongly fit, 29% mostly fit, 35% fit are included) thought that the system can effectively help their learning. When were questioned subjectively pleasing, only 9% stated that they are not willing to use the system, which indicates success of the system. However, 28% of the users rarely contribute to the terminology system, including modifying or adding term. This shows the limitation of arousing participation interest. And the last question indicates the interaction among the users should improve in future.

6 Conclusion and Future Work

We have presented an integrated terminology system that supported by Ajax technology and web-based development systems, targeting at the help for learner to conveniently find the relevant information when he encounters a new term in e-learning. The system includes four main modules, term display module, term query module, term modify module, term add module, which can provide immediate assistance to learner when he encounters a new term and tries to find the explanation of term. In addition, the system uses Ajax to increase user experience in comparison with the traditional terminology system. Exploiting with Ajax, the system takes the following measures to provide student with a student-centered e-learning environment, automatically add explanation to term, search suggest, save automatically, preload page etc.

We made an evaluation on the system in the section 6. In future, we will improve its performance to provide the better help to the learners.

Acknowledgements

This work is partially supported by the Natural Science Foundation of China under grant Number 10771120, Key Project of Culture Department of Shandong Province and Instructional Innovation Project of Education Department of Shandong Province under Grant NO.B05042; Project of Outstanding backbone University teacher international cooperative Cultivation of Shandong Province, P.R.China.

The authors would like to give their sincere thanks to Prof. Shumin Kang, Prof. Siwei Bi, Dr. Yuxia Lei, Dr. Zili Zhou, Dr. Xiulan Li and other colleagues and friends for the pleasant cooperation in Project of Instructional Innovation of Education Department of Shandong Province and also for their constructive proposals.

References

1. Chinese Computer Terminology System, `http://ccts.cs.cuhk.edu.hk/`
2. The basketball dictionary of the hoopchina, http://www.hoopchina.com/ext/index.php/
3. Horton, W., Taylor, L., Ignacio, A., Hoft, N.L.: The Web page Design Cookbook. John Wiley, New York (1996)

4. Downes, S.: E-Learning 2.0, http://www.downes.ca/post/31741
5. Karrer,T.: Understanding E-Learning 2.0, http://www.learningcircuits.org/2007/0707karrer.html
6. Deng, G.: Pilot Study of E-learning2.0 and its Teaching Models, http://www.etc.edu.cn/show/2007/elearning09.htm
7. Novak, J., Gowin, D.: Learning How to Learn. Cambridge University Press, New York (1984)
8. Garrett, J. J.: Ajax: A New Approach to Web Applications, http://www.adaptivepath.com/ideas/essays/archives/000385.php
9. Wang, P., Fang, M.: Conquer Ajax: detailed explanation of development technology of web 2.0 (2006)
10. Resource centre of Ajax technology, http://www.ibm.com/developerworks/cn/AJAX/
11. Whats AJAX, http://AJAXpatterns.org/wiki/index.php?title=Whats_AJAX
12. McLaughlin, B.: Mastering Ajax, http://www.ibm.com/developerworks/xml/library/wa-ajaxintro1.html

Idea and Practice for Paperless Education

Yiming Chen and Lianghai Wu

Maoming University, Maoming 525000, Guangdong, China
gdmmcym@126.com, webisland@126.com

Abstract. This article introduces the concept and style of E-Learning, analyzes the Blog's characteristic while being applied in teaching, then designs an E-Learning platform model based on Blog. At last, it also discusses the model's function, characteristic and existing practice problems.

Keywords: Blog, E-learning, Model, Collaborative learning, resource management.

E-Learning is a new type of learning based on the Web. Blog is regarded as a new web culture, and the application in education and instruction is a hot point in recent research. We have done some research and practice on how to merge Blog into E-Learning.

1 Bolg and Its Instructional Features

1.1 The Introduction of Blog

Blog (Weblog) is the serial accounts or diaries on the Web. It is a personal website with its content given by the owner, listed from the latest. It has a social function. It appears in the form of text and super-media. The users simply publish their individual knowledge, ideas, thoughts and news as posts to share on the web. Blogs integrate articles, accounts and communication together to make them into a web-page, with the information much larger than traditional diaries. The blogs'information is filed by date, classification and tag.

1.2 The Instructional Feature of Blog

Blogs can be used as individual or course resource, and as exhibitive center and communication center in the process of education and instruction with the features as follows:

(1) In the process of instruction, blogs can be organized in the form of "basic blog"and "group blog". Basic blog is the simplest form with teaches or students supplying certain topic and relative resources for discussion or comment. Group blog is the transformation of basic blogs. It consists of groups of students to complete the blog diaries together. The authors can edit not only their own diaries but also others'in the same group, which permit them to discuss on the same topic and even complete a collaborate project.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 147–152, 2008.

(2) Blog is a constructive learning tool suitable for the students'cognitive psychology. It can organically integrate "situation", "collaboration", "dialogue"and "meaning construction". The value of blog in instruction can be seen in the following aspects:

- expanding and extending the space for teaches and students to communicate;
- fostering the interest of self-seeking, and promoting collaborative level; recording the whole learning process, and systematically realizing knowledge management;
- accounting and evaluating together, and recording the feeling, discussion, reflection in the learning process to make a all-round evaluation.

(3) Blog, as a new form of communication on the web, will help the students to enhance their abilities in receiving, analyzing, and processing information when it is used in instruction.

Blog is now being dynamically growing up. We can see easily that blog changes knowledge, news, and learning. It has great potentiality in the process of changing the world.

2 E-Learning and Its Learning Formation

E-Learning, in its broad sense, refers to the learning through electronic media with three general forms: Learning through satellites'TV system, Learning through audio-video meeting system, and Learning through computer web system. The narrow sense of E-Learning refers to the learning on-line or web based learning. It is simple the learning through computer network with the web learning environment consisting of multimedia web learning resources, on-line learning community, and web technical platform.

E-Learning has synchronous and asynchronous forms.

(1) Synchronous E-Learning refers to the learning in virtual classrooms. The instructional information is shown as text or projection on the presentation device. Sometimes the recording of the teaching can be presented directly on the line. Powerful device and technical support is needed in this form of learning, to guarantee the fluent communication between teachers-students and students-students. The advantage is that the teachers can provide simultaneous feedback to students and the simultaneous communication can make a real classroom environment, which gives the students a feeling of family-ship. The disadvantage is that a formal timetable is needed, which makes it not flexible enough.

(2) In asynchronous E-Learning, learners can make their own pace of web course learning. They can check their learning record at any time and learn by their own learning and thinking habit. They can also get the personal feedback from the instructor. Students can communicate with their classmates and teachers on-line, exchange e-mails, have class or group discussion through the web. This form of learning is flexible and easily self-controlled.

3 The Model of E-Learning Platform Based on Blog

Blog can be regarded as an instructional tool, as well as a learning model, while E-Learning a brand-new learning method based on the web. The application of blog used in instruction can be generalized as automatic, (inter) collaborative, and individualized. It can merge with E-Learning to form an instructional model based on blog to provide a learning and communication platform for learners and instructors. Figure 1 shows the model.

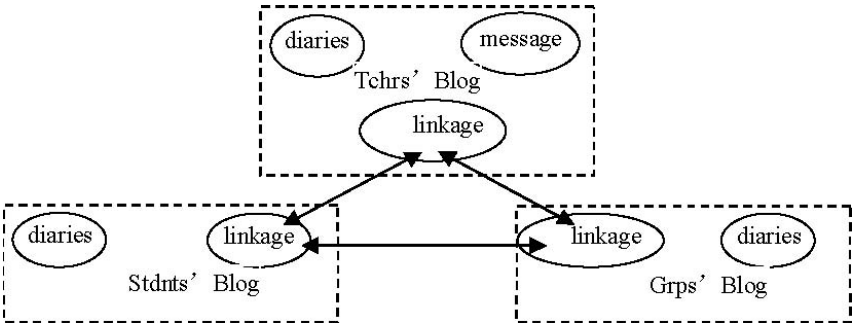


Fig. 1. The model of the platform

3.1 The Connotation of the Elements and Device Function

The model is a loose colony of teachers'blog, students'blog and groups'blog. The teachers'blog is the center to organize the students'active learning, forming the platform of synchronous and asynchronous learning.

Teachers are the principal part of the model with three functional sections, including diaries which present the introduction, task, process and resource of the topic with text, pictures, multimedia courseware and web-page, message which is the window for teachers to communicate with the students, and linkage which implements the link of blogs between teachers and students to support the share knowledge, search and inquiry, and find the relative information quickly. The function of the teachers'blog(Tchrs'Blog) is to provide tools for self-improvement, knowledge management, and a platform for communication, and an assistant for instructional research. Figure 2 is the interface for the teachers'blog.

The students'blog (Stdnts'Blog) and the groups'blog (Grps'Blog) are the basic parts of the model. The students can use the blogs to do any topic-round activities, to collect relative web resources, to publish topic diaries (reflection, E-text, activity accounts etc.), to read the latest diaries of others', and take part in the comment, self-evaluation and evaluation of the diaries.

3.2 The Characteristics of the Model

(1) The full advantage of the space on the web is used for effective E-Learning. Firstly, the teachers and students can choose and register their free blogs at many web sites

Fig. 2. The interface for the teachers'blog

providing blog spaces. Secondly, Each part on the platform is equal at the inter- or multi-communication.

The teacher is the center of the model. He/She adds the URL of the students to his/her blog in the linkage of the classified catalogue to examine, supervise and evaluate the learning process of the students. The students, at the same time, will answer the questions or hand in their assignment in accordance to the teacher's requirement in the blog. The students can choose to read the blogs and give comments as a passer-by. They can record their acquisition and experience in their blogs. The blogs digitalize the files and notes of learning and record the behavior of the students, such as their feedback to the instruction, their report to social practice, and the collection of their learning resources for investigative learning. The platform, on the other hand, aims mainly at the process not the results of the learning, where the views of the teachers or the books are presented. The students can publish their understanding to them here, and the teachers can comment or provide a guidance to students'questions and make the process into a dynamic interaction between teachers and students.

(2) The learning resource is digitally managed. Learning resource is the sustainable guarantee for the instructional activities on the platform. The management of the learning resources is an important task in the model of the platform. The model supports the upload and download of various multi-media files, which can add and update the learning resource constantly. The content in the blogs includes reference materials, courseware, teacher plan, experiment instructions, works of students, discussion on important and difficult points, and exercises. The research of the teachers to the course and production, the papers, and the honor the teachers have received can also be included. Figure 3 shows the functional framework of resource management.

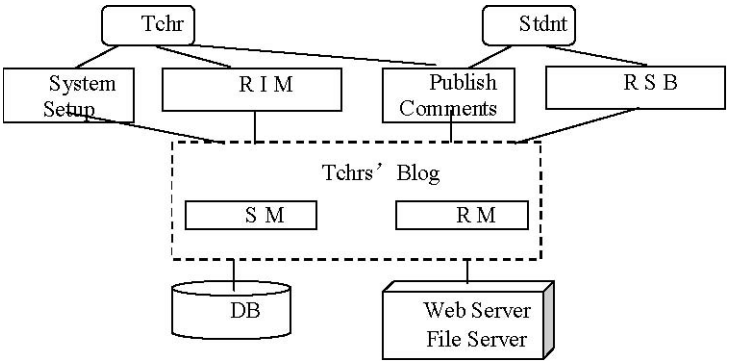


Fig. 3. The functional framework of resource management of the Blog. RIM : Resource Issue and Management; RSB: Resource Search and Browse; SM: System Management; RM: Resource Management.

(3) All-round evaluation can be implemented to the learning process. The evaluation of the instructional effect can be implemented with three sections ---individual, group(in-or inter group) and teachers. Process and qualitative evaluation are recognized in the platform with importance attached to comprehensive evaluation to the learners. The evaluation can be done simultaneously on the web by the reflection to detect the redeeming feature and the deficiency, and to find new problems arousing new elicitation. The teacher occupies the dominant position in the instructional process. He/She organizes the instructional activities, handles the feedbacks and evaluate the learning. The students communicate and feed back in the blog when the blog record their learning developments. The record in the blog shows the behavior and contribution of the students, supplying reference for the evaluation.

(4) Self-determined, collaborative and individualized learning can be achieved. The platform based on the web provides a effective means for the learners in different places. They can learn synchronously and asynchronously in a virtual classroom with the technology of the web, database and artificial intelligence.

3.3 The Limitation of the Model

(1) The model is public and incompact. It can act only as a complement tool but not a typical criterion for education and instruction management.

(2) The ideas of liberalization and individualization may drop a hint for the students to publish irresponsible comments or do improper operation.

(3) The resource management has only one-level classification, which may cause indistinct navigation. Only full-text search can be done with no classified search, and may be inconvenient for searching the resources. The resources are appended with texts, and the labels of attribution are not criterion.

4 Conclusion

Blog is an instructional tool or model based on the web. It is individualized, interactive and convenient for knowledge management. E-Learning is a fashion for learning

on-line, which makes the learning automatic, interactive, collaborative and individualized. The research and practice we have made aim at the amalgamation of the two to compose a model and provide a platform for the learning and communication of teachers and students.

References

1. Cuifeng, Z.: E-Learning:Product of Information Age. Vista of Global Education (5), 52–53 (2002)
2. Lianghai, W.: Blog's Appliction in Learning Source Management. China Science and Technology Information (21), 248–249 (2007)
3. Ming, Z., Youfu, D.: Construct the Cooperation Learning Platform with Blog. Journal of Educational Technology (4), 4–5 (2006)
4. Qing, W.: Construct the Remote Education Teaching Mode Integrating Blog with Webquest. Journal of Nanjing Radio & Television University (1), 13–16 (2006)

SyTroN: Virtual Desk for Collaborative, Tele-operated and Tele-learning System with Real Devices

R. Chellali[1], N. Mollet[2], C. Dumas[3], and G. Subileau[4]

[1] IIT Genova, Italy
ryad.chellali@iit.it
[2] IIT Genova, Italy
nicolas.mollet@gmail.com
[3] IRCCyN Nantes, France
cedric.dumas@emn.fr
[4] Virtools/Dassault systems Paris, France
geoffroy.subileau@virtools.com

Abstract. Tele-training is a main issue nowadays and it is strongly motivated by the increased mobility of people. This mobility shouldn't be a limit for a good training: pertinence and efficiency have to be the core of distant training environments. In this paper, we introduce SyTroN, a tele-learning system using virtual reality and tele-operation techniques. The first aim is to propose intuitive virtual classrooms/desks supervised by a real teacher for collaborative or individual distant learning, using Internet. The second goal is to go from virtual to real: SyTroN supports the connection to real devices, potentially rare and expensive, allowing distant experimentation's abstracted by virtual tools. After 5 years of development, our work has been validated with psychologic tests, which highlight the efficiency of our global system on one year of use within an engineering school.

1 Introduction

The use of virtual reality has been strongly facilitated during the last ten years: techniques are mature, costs have been reduced and computers and devices are enough powerful for real-time interactions with realistic environments. Industrialists have also validated huge prototyping, simulating and training systems: we can cite the VTT project, which aims to train on technical gestures while manipulating a virtual milling machine[1]. We can also cite the GVT project[2], with individual and collaborative [3] learning for industrial maintenance procedure on military equipment. The use of virtual reality for large public is more recent, but already pertinent: interactive 3D video-games using VR devices, such as the Nintendo Wii, is a good example of this success. On another hand, many works have been already done in the field of distant/e-learning, illustrated for example by the notion of virtual classrooms/courses proposed by some universities[4] with pertinent results[5]. Those distant learning are now very useful in a world where people are traveling a lot, and where jobs opportunities conducted them to move several times. This context makes (distant and mobile)

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 153–161, 2008.

training as one of the most active domain in the field of Virtual Reality. Now, techniques could go from expensive training environments in dedicated rooms to personal computers or even laptops, while large-band networks allow easy access to distant and heavy information.

Concerning Tele-operated systems, virtual reality techniques are very interesting and used since the begins. Indeed it provide a good abstraction for real devices operators have to operate with. VR allows simulations, but also interesting augmented representations of real devices and this is very useful in executing tele-operated tasks. Obviously, one can take advantage of these rich environments for pedagogical purposes. We can cite the work of Bicchi[6], who applied e-learning techniques in the field of automatic control, in order to learn the usage of a robotic arm for example[7]. Dongsik[8] was focused on simulating electronic circuits in distant virtual laboratory, with the ability to apply the models on real equipment with webcams, in order to validate the theoretical simulations.

SyTroN project is born in this context. Fundamental research and innovative solutions have been developed in order to success in mixing VR techniques, mobile and distant learning, and tele-operation applications. With those techniques, the fundamental goal of our project is to increase the learning efficiency in a two steps process: first, students learn and simulate, and then they manipulate real devices. To do so, SyTroN provide virtual classrooms supervised by a teacher, where distant students, with standard supports such as distant reference books or media contents, have to train by acquiring theoretical skills and learn more about processes and devices models. Once an acceptable level reached, students can move to teleoperation of the real devices to face and solve real problems. SyTroN system is a complete and functional solution allowing for the moment simulation and training on 3 different tele-operated devices. The system also takes advantage of VR for pedagogical add-ons not available in the real world, in particular with dynamical and contextual information in the 3D environment to help the training process.

The first part of this paper deals with the presentation of our models and its implementation. The next section gives an overview of the system usage. Finally, the last section presents the field validation of SyTroN, based on a comparison of training sessions with traditional techniques versus training sessions using SyTroN.

2 Contribution: SyTroN

We present in this section the models and the implementation of SyTroN.

2.1 Global Vision

The figure1 illustrate the global vision of SyTroN. The proposed architecture allow multiple distant connections of all the elements of the system. We can identify three logical sub-systems, distributed on the Internet network:

– The users, which include the teacher and the students.
– The devices
– The manager and the knowledge database
– The devices
– The manager and the knowledge database

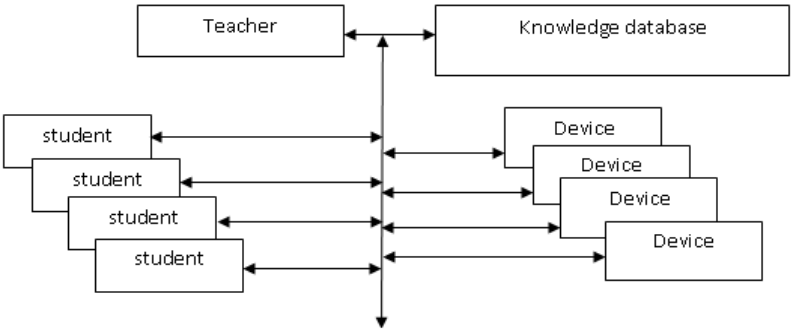


Fig. 1. Global vision of SyTroN

Figure 1 : Architecture of SyTroN

The teacher and the student sub-systems are user-oriented, i.e. the interfaces are designed to ease the use and to perform the learning process. The devices entities are control oriented, i.e., each entity achieves mainly the functions of enabling the remote control of any device of the platform. A device is composed itself of two parts: the device and the device server. The manager is the core of the system. It contains all the information describing the system (users, devices and network), like status, configurations and history for instance. It is database oriented.

Here's an overview of the standard *usage* of the system:

– *Teacher connection:* a teacher requests a connection. Verifying his rights from the knowledge database does this. Thus, the teacher is given the list of actions he is able to do. These actions are of two types: time-critical or asynchronous.
– *Student connection*: the student can connect using the same protocol as teachers. Obviously his rights are different. Once connected, the student can work alone or within an existing virtual classroom. For both modes, the student accesses the following services.
– *Getting or uploading a document:* before beginning the lecture, the student will download the contents of the course to study the basics. Notice that this is possible if the teacher has already uploaded the needed material.
– *Discussing with the teacher:* during a learning session or off-line the session, a student can contact a teacher or the other students to have a chat, a videoconference or just by sending an email. Here also, two categories of functions are available, synchronous or asynchronous.
– *Using a device:* a user can interact with a physical device in real time. This is the main sensitive action a user can perform. Following the availability of the device and the user rights, a device is allowed to one user (except when the device is shared between the teacher and the student). The interaction here is mainly synchronous and time critical. Indeed, the control loop is not local but it is geographically distributed over the network.

2.2 Functions

In the following points, we are now introducing the set of functions attached to each logical entity, according to the needs of usage.

Users functions. Except for some specific functions, students and teachers are considered as end-users and both are clients. The corresponding functions enable to access the SyTroN services, namely, the device servers and the knowledge database. Two main sets of functions are available: Communications, and Remote controls.

The *communications* functions enables the users to access to the contents and to people within the SyTroN system. Namely, the students can:

– Access to lectures, tests and evaluations ;
– Contact teachers and other students ;
– Get the devices status and gateways ;
– Simulate devices;

Some specific functions are reserved to teachers : adding devices, adding lectures, and setting tests and evaluations.

The *remote control* functions link users to physical devices. Users are enabled to achieve remote physical interactions: they may send controls and receive measurements in return, i.e. the device status as well as some information about the remote environment.

The previous functions are implemented for the most used OSs: Windows, Unix and Linux. Users terminals are designed for WEB-based interface as well as proprietary interface. The communication blocks of each function are TCP-IP based.

Device functions are separated between a server and a controller.

Device server is the gateway between users and the remote world. It is composed of two parts: the device gateway and the device controller. The device gateway connects the users and the knowledge database to the device. The device controller pilots the actuators and gets the sensors information. The device gateway is just a translator. Depending on the considered device, the gateway formats the users commands and passes the parameters to the controller. The concerned actuators execute the command and respond to the device controller by giving the new position of the mobile. The functions of the gateway are mainly dedicated to synchronous communication. TCP protocol is used to guaranty the exchanges synchronicity and integrity.

Device controller is the interface between the real and the logical worlds. Its main function is to close the local control loop. This component is device dependent: for each device, one needs to write a specific driver. For SyTroN implementation, three devices with three different dynamics have been tested. Due to the nature of the considered devices, two control platform are used : the Matlab (from MathWorks) platform and a proprietary platform (C++ based API).

Knowledge database functions. The knowledge database contains all the information needed to manage the global system. Three sub-databases describes the three logical units (the teacher, the student and the device) involved in a distant learning session, the cross-relationships between the units, and a history of platform uses.

The main functions of the K-database are the following:

– Managing people by setting and verifying the rights of users such as name, co-ordinates (email, address), list of lectures (etc..). Each field allowing to users to access to SyTroN services.
– Managing contents such as Lectures, Tests and evaluation program, Tests and evaluation results.
– Managing devices by adding, modifying or keeping devices histories.

This database is SQL compliant. The current version is MySQL and specific C++ API is used to interact with users and devices. In addition, we added a gateway module to open partly the system to let administration manage the teaching activities/scheduling within the engineering school.

3 Example of Devices and Usage

We give in this section a short description of two devices we have integrated in SyTroN, and some views of the running system.

3.1 Devices

The figure 2 presents two of the three devices we have already included in our system.

The heating board is dedicated to the study of heating processes, simulating complex systems with numerous inputs and outputs in the state space (an algebraic approach of automatic control). The heating board is a basic MIMO system (Multiple Inputs Multiple Outputs). The students have here to keep constant the temperature of the board, regardless to the external temperature. The goal of the training session is to learn MIMO control techniques, especially control laws.

The mobile Pekee robot is a classical robot produced by Wany Robotics SA. This robot has two actuators, namely two DC motors enabling the robot to move with 3 degrees of freedom. On the other hand, Pekee has onboard a set of sensors (telemetry, light, gyros, collision detector, etc.) enabling to handle the robot environment. Both actuators and sensors are handled by the onboard PC. This last communicate with any other PC using the wireless Ethernet based channel. The communication can concern the state of the robot (sensors information) or robot controls (DC motors or a request for a specific measurement).

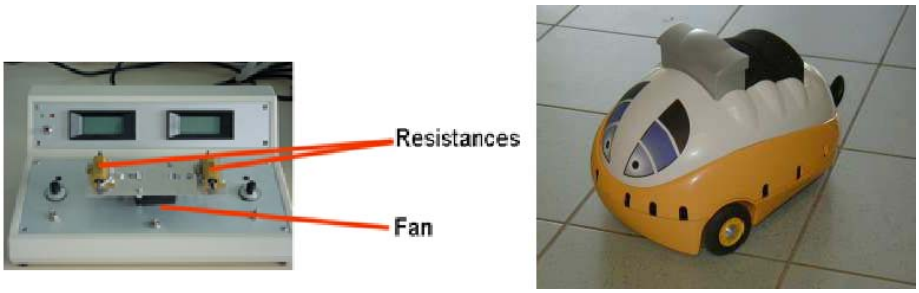


Fig. 2. -*Left*-The heating board. -*Right*-The Pekee robot.

3.2 Usage

Starting the session, the student download the documents explaining the lecture, the manipulation protocol and the attended results. Following this, the simulation phase can start on the virtual device (equivalent of the real one). This enables to acquire the theoretical basis. In this virtual classroom(figure 3), students can easily manipulate virtual devices, helped by VR pedagogical metaphors such as avatars, the use of transparency, etc.

The heating board. The student can verify the results of the simulations by comparing resulting curves to theoretical ones. From the derived model, the student can switch to the real device to setup the control parameters, namely the PID constants. These parameters are sent via the asynchronous TCP channel to the device server and the real test can start. The control closed loop is then operating to maintain the temperature to a desired one. The client calculate the inputs of the heating board and send it via the TCP synchronous channel. The device extract the current temperature and return it to the client and so on. This question-response process is done at 1Hz frequency. Indeed, as the response time of the system is greater than 15secondes, 1Hz frequency for the control loop is enough to verify the Shanon theorem and to ensure the stability of the system. To add more impact on the learner, a video feedback was added to this device.

The mobile Pekee robot. For the mobile robot, time constraints are stronger than for the previous device. Indeed, the purpose here is to enable to a distant student to pilot a mobile with a visual and a force feedback. It is obvious that the shorter the system response time is the better is the manipulation. A UDP based protocol is used for this service. As for the heating board, the first stages of the mobile robot manipulations are concerned with the discovery of the device.

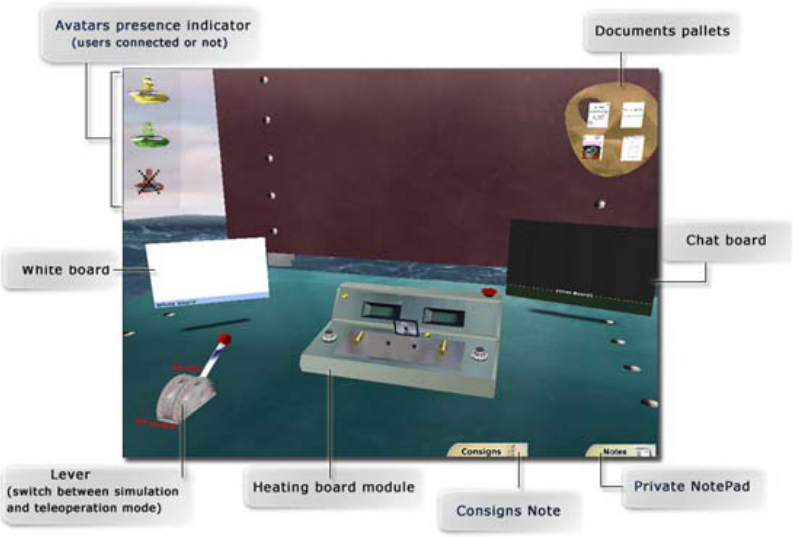


Fig. 3. The 3D desk and its components

This is done using documents and simulations. For the tele-operated mode, the user may use a force feedback joystick, a mouse or the computer keyboard to pilot the robot. This generate motion commands that are sent to the robot. The robot execute these commands, capture about 31 sensors measurements and reply to the client. These measurements are then used to refresh the client virtual environment. As the robot speed is about 0.5m/s, the user may react in less than 1/2 seconds to avoid obstacles in front of the robot. This response time is obtained by tuning the range sensor dynamics to 1m depth. Following that and considering transmissions time delays in the network, any object about 0.25m has to be considered as very close obstacle. A hidden mobile robot was added to the interface to take into account the effects of time delays and the resulting offset between the user command and the real state of the remote robot.

The classroom metaphor : sharing the working environment. One of the main issues of Sytron is to offer a collaborative platform. Collaborative services use the network to share users activities to facilitate the realization of tasks and communication, by displaying graphical feedback to show connected people and what they are doing. For Sytron, the collaborative framework and the share of the 3D virtual environment are built around a communication server named reflector. The Reflector supports a set of Peer to Peer or broadcast-like services (chat or voice conferencing for instance). UDP and TCP based protocols are used to handle these services with no specific constraints except the users comfort.

The virtual white board is designed to let user share results and formula. With this tool, users can draw whatever they want thanks to the pen tablet (or mouse). The chat board display messages written by users who can type a message at any time. Users can also see the avatars corresponding to the other users connected. When a user has a service as focus, his correspondent avatar flies next to the object corresponding to this service, in order to indicate to the other participants what he's doing.

4 Evaluation

Before developing and writing codes for Sytron, we started by designing a paper prototype based on interviews we done with all actors. Teachers were asked to detail their presence-based lectures regarding structural aspects, timing, exchanges (teacher-students, students-students) during lectures, the most impacting parts, etc. This study leads us to write a script with handmade screens and concepts to include into the virtual desk and the virtual classroom. Our approach was not only objective: some of our considerations are purely subjective to take into account personalities of teachers and how they transmit knowledge. Students also evaluate the quality of knowledge transmission mostly perceptually. They are more sensitive to the way the teacher introduces and explains concepts than the validity of concepts. Following that, we opened the system to two categories of students: normal program for young engineer (12 people) and adult program for which students work mainly from home and nightly (10 people). Individual usages were checked and we verified our pre-conception hypothesis: the interface is fluid and users discovered services in a quasi natural way. The second verification was concerned with the shared space and attached tools. Adults and young engineer reactions were different: for the first category, chat and live talk were more used before and during experiments. For young people exchanges were mainly concerned with simulations/real

experiments results. May be cultural gap exists between the two categories and ones are more used with known channels (MSN or Skype for instance) and our offer was disturbing their communication environment.

The last evaluation to do is concerned with the classroom metaphor. All participants may use the system at the same time with the goal to replace the existing lectures. This step is not of evidence and we have to reconsider the previous evaluation approach: marks, credits and public evaluations are to be produced.

5 Conclusion

We detailed the Sytron system, its architecture, its components, its use and some preliminary evaluations. First, we strongly concentrate our work on automatic control aspects (closing high/low speed loops over Internet) which are not described here. The second step was to pre-design the virtual desk and its components. Voice, video and written messages channels were integrated within the system. A central system managing technical and administrative matters finishes our core developments. We have an operational system which can constitute a good basis for further improvements like adding an ITS intelligent Tutoring System for automatic assistance for the teacher. Indeed, teachers currently act in an old fashion by following learners. An ITS can help to personalize more the relationship between each partner. The other contribution we made and which must be also improved in the future is the use of real devices to let learners confront to real problems. In addition VR techniques enable us to augment the reality by adding guides and milestones. This let users use the maximum available information at experimenting time.

Special acknowledgments to Amelie Imafouo. SyTroN can also be tested at http://www.emn.fr/x-auto/reposit .

References

[1] Crison, F., Lecuyer, A., Mellet dHuart, D., Michel, G., Burkhardt, J.M., Dautin, J.L.: Virtual technical trainer: Training to use milling machine with multi-sensory feedback in virtual reality. In: IEEE VR 2005 (2005)

[2] Gerbaud, S., Mollet, N., Ganier, F., Arnaldi, B., Tisseau, J.: Gvt: a platform to create virtual environments for procedural training. In: IEEE VR 2008 (2008)

[3] Gerbaud, S., Mollet, N., Arnaldi, B.: Virtual Environments for Training: From Individual Learning to Collaboration with Humanoids. In: Hui, K.-c., Pan, Z., Chung, R.C.-k., Wang, C.C.L., Jin, X., Göbel, S., Li, E.C.-L. (eds.) EDUTAINMENT 2007. LNCS, vol. 4469, pp. 116–127. Springer, Heidelberg (2007)

[4] Lepori, B., Cantoni, L., Succi, C.: The introduction of e-learning in european universities: models and strategies. In: Kerres, M., Voss (eds.) Digitaler Campus, pp. 74–83 (2003)

[5] Kukk, V.: Analysis of experience: fully web-based introductory course in electrical engineering. In: Proc. Int. Workshop e-Learning and Virtual and Remote Laboratories, pp. 111–118 (2004)

[6] Balestrino, A., Bicchi, A., Caiti, A., Calabr, V., Cecchini, T., Coppelli, A., Pallottino, L., Tonietti, G.: From tele-laboratory to e-learning in automation curricula at the university of pisa. In: IFAC 2005 (2005)
[7] Bicchi, A., Caiti, A., Pallottino, L., Tonietti, G.: On-line robotic experiments for tele-education at the university of pisa. International Journal of Robotic Systems (2004)
[8] Kim, D., Choi, K., Lee, S., Jeon, C., Yoo, J.: Implementation of a web-based hybrid educational system for enhancing learning efficiency of engi-neering experiments. In: Technologies for E-Learning and Digital Entertainment -Edutainment 2007. LNCS, pp. 357–368 (2007)

An Examination of Students' Perception of Blended E-Learning in Chinese Higher Education

Jianhua Zhao

School of Information Technology in Education, South China Normal University,Guangzhou 510631, China
jhuazhao@gmail.com

Abstract. Blended e-learning is a brand-new approach which is used for technology integration currently in China. Meanwhile, many kinds of Virtual Learning Environments (VLEs), such as WebCT, BlackBoard, and WebCL are emerging in the field of Chinese Higher Education in the past decade for serving this new model. How do Chinese students think about this new approach? There is no much research addressed on it. We design a small-scale research for examining students' perceptions of blended learning, which is based upon an undergraduate course offered at School of Information Technology in Education in South China Normal University. 20 students registered in the course. The findings of this study reveal that students' understandings of blended learning are formed through the course, and their attitudes toward it are positive.

Keywords: Blended learning, online learning, face-to-face learning, virtual learning environment.

1 Introduction

Blended learning is hardly to say that it is a brand-new term, rather "an old friend gets a new name" [1], which can be considered as a fairly new term in education lingo, but the concept has been around for decades. Blended e-learning offers a new learning approach for combining different delivery modes, normally is online and face-to-face teaching [2][3]. It changes the traditional face-to-face learning through bringing a non-geographical and non-time limitation online learning when it is applied in on-campus teaching/learning, because teachers and students, students and student can meet online when they are off-campus. Blended learning is the fastest growing trend in e-learning [4], and the successful e-learning program is or will become a blended learning program [5]. Therefore, blended learning could be the next most popular term which can replace e-learning. This is the reason for using blended e-learning rather than blended learning in this study.

Numerous researchers have presented their efforts on exploring the meaning of blended learning. For example, Smith defines blended learning as a method of education at a distance that uses technology (high-tech, such as television and the Internet or low-tech, such as voice mail or conference calls) combined with traditional (or, stand-up) education or training [1]. Bielawski and Metcalf point out that blended

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 162–170, 2008.

learning focuses on optimising achievement of learning objectives by applying the "right" learning technologies to match the "right" personal learning style to transfer the "right" skills to the "right" person at the "right" time [6]. Ward and LaBranche consider blended learning as learning events that combine aspects of online and face-to-face instruction [4]. Even researchers have diverse perspectives, blended e-learning should combine online and face-to-face instruction, well-designed and optimised learning delivery methods, and specific objectives, learning events, tasks, and purposes are involved in a blended e-learning process.

There is only a short period since blended e-learning has been introduced and utilised in the field Chinese higher education [7][8]. How Chinese students think about this brand-new approach? There is no much research addressed on it. The purpose of this study is to examine students' perceptions of blended e-learning in the setting of Chinese higher education. The field studies have identified the causal relationship between one's perception and one's performance. Such as Ginns and Ellis examine the quality of blended learning and the results show that the approaches students take to learning, and the subsequent quality of their learning is closely related to their perceptions of their learning experience [9]. Delialioglu and Yildirim investigate students' perception of the "effective dimensions of interactive learning" in a hybrid course, and the finding demonstrates that the way instructivist and constructivist elements are blended, the need for metacognitive support, authentic activities, collaboration, type and source of motivation, individualized learning, and access to the internet played important roles in students' learning in the hybrid course [10]. Reichlmayr claims that the use of blended learning techniques helps to make the experience a more satisfying one by increasing the effectiveness of collaboration in team activities and interaction with the instructor through the application of distance learning and social computing technologies [11].

A small-scale research is designed to examine students' perception of blended e-learning and it will open a door for understanding students' thinking and behaviour in the blended e-learning environment in Chinese higher education. Three research questions are developed for this study, which include:

- How do the students perceive blended learning?
- What do student experience classroom-based learning in a blended learning environment?
- What are students' experiences of online learning?

2 Method

Questionnaire survey research was used as method in this study. In order to compare students' perceptions of blended e-learning before and after the course, the pre- and post-surveys were designed for collecting the relevant data. Except the closed questions, the opened questions are involved in the questionnaire too in order to in-depth examine students' thinking and understanding.

20 undergraduate students at the School of Information Technology in Education (SITE), South China Normal University (SCNU) registered in a bilingual course – Computer in Education for eight-week studying. 20 students participated in the first questionnaire survey at the beginning of the course. The responded rate and the valid rate are all 100%. 16 students participated in the second questionnaire survey at the end

of the course, and 13 of them gave their response back. The responded rate and the valid rate are 81.25%.

The learning environment in this study includes two settings, which one is ICT-based classroom with internet connection, and another is a VLE – WebCL which is developed by Beijing Normal University. This is a good environment where it offers lots of functions, such as group and class discussion forums, course management module, group management module, and personal blogs etc.

3 Data Analysis

3.1 Students' Experiences of Blended E-Learning

With respect to the first survey of this study, the question "do you have any experiences of blended e-learning" was proposed to find out whether they had experiences or not (see Table 1).

Table 1. Students' experiences of blended learning

Examined items	The status of students using the blended learning	
	Frequency	Percentage
Used it before	4	20
Never used it	16	80
Heard this term before	5	25
Never heard this term	15	75

The data demonstrate only 20% students chose "used blended learning before", and 80% students chose "never used it". The data also illustrate 25% students heard the term "blended learning" before, and 75% students "never heard this term". The data reveal that most continuing students were lack of the experiences of blended learning at the beginning of this course.

The further analysis of students' choice in questionnaire survey, when students were asked to give "where you heard this term" if they chose "yes, I have used it before", they all replied "from this course" (100%). It is an interesting answer because it reveals that students have no experience of blended e-learning before this course.

The qualitative data collected from the second questionnaire survey were analyzed in order to examine how their experiences of blended e-learning developed through the course.

> ***Miriam:*** *Blended e-learning combines the conventional educational approach (classroom-based education) with the new educational approach (i.e. web-based education) together. These two approaches can be complementary each other and the new approach represents their advantages from that two approaches. It can improve the quality and effectiveness of education.*

Miss Mariam's thinking was similar to the introduction from the relevant literature [12][13], which demonstrates that she already formed appropriate understanding of blended e-learning. Her attitude to blended learning was positive.

> ***Onjcn:*** *Blended learning combines online learning and FTF learning. Online learning can improve the learning efficiency and to facilitate students learning. In-depth understanding can be constructed through the communication among students though the web. Meanwhile, students can consider the understanding from the classroom as the foundation for further learning. Online learning can be used to in-depth explore the relevant knowledge. FTF in the classroom and online discussion are all very useful for improving their understanding.*

Mr. Onjcn introduced his understanding of blended e-learning from online learning and FTF learning. He explained how these two learning approaches were useful for their understanding. He considered communication as the essential element. It could be acquainted from classroom-based learning first and then was used as the foundation in the online learning environment. The discussion was a useful means for building students' understanding.

> ***Tina:*** *In blended learning, tutor can give her/is guide, lecturing, and explanation to students. It can improve the efficiency of the class. Web-based learning resources can supply the shortage of the resources in the classroom setting. Therefore, students have more chances to learn the relevant information.*

Miss Tina explained the reason why students had more chances to acquaint more information, because the online learning resources could remedy the shortage of the resources in the classroom-based setting. Normally, it was difficult to get sufficient learning resources in the classroom-based setting. However, when students were in the online-based setting, they could be supplied plentiful learning resources. Blended e-learning provided chance to compliment the shortage of classroom-based learning.

> ***Gigi:*** *In a blended learning environment, classroom-based and online-based instruction can complement each other. Online course can assist classroom-based instruction. There would be more time to collaborative learning between students and tutor, students and students.*

Miss Gigi described the relationship between tutor and students, students and students in blended e-learning. They would have much time for collaborative learning, because online courses could assist their learning. Online learning could help students to know more background of the relevant knowledge in their courses. Meanwhile, they could also learn from online course themselves. It was useful to improve the efficiency and effectiveness of collaborative learning when students and tutor were getting together.

3.2 Students' Experiences of Classroom-Based Learning

As a blended e-learning environment, the classroom-based learning environment should be effective as well. In the second survey (anonym) of this study, students were required to give their opinions, suggestions, or comments to the question "what kind of classroom-based environment is effective". Some of their descriptions are quoted and analysed as follows:

> ***Student A (anonym):*** *Interactive teaching approach should be increased in the traditional classroom. Knowledge infusing should be changed in terms of the learning contents. Heuristic is preferred in the conventional teaching environments.*

We do not like spoon-feeding teaching. Teaching techniques should be improved and multimedia, networked technology should be used in the class.

This student suggested that the interactive teaching approach should be used in the classroom-based approach. Heuristics was the priority method in the conventional teaching environment. Student did not like spoon-feeding teaching approach at all. This student also suggested using ICT in the classroom, which represented most students' considerations. The conventional teaching approach should be changed in an e-learning environment. Actually, many tutors were trying the new student-oriented approach in Chinese educational field, such as constructivist learning approach, group learning, or problem-based learning [14].

Student B (anonym): *Teaching activities should be added some vivacious contents. Their purposes are to attract students' interest.*

The suggestion from student B reveals that they did not like the bald or duck-feeding approach. Teaching activities should be vivacious, interesting, and attractive. What this student introduced was related to the techniques of the lecturing in the classroom.

Students' considerations on the effective classroom-based learning setting are categorized and presented in Table 2.

Table 2. Students' consideration on the effective classroom-based learning

Consideration	Frequency	Percentage (%)
Interaction	6	31.6
Learning Environment	9	47.4
Engagement Learning	10	52.6
ICT Application	7	36.8
Teaching Techniques	9	47.4
Heuristics Teaching	8	42.1
Group learning	14	73.7

The data in table 2 reveal that "group learning" was the most important element for the effective classroom-based learning, because fourteen students (73.7%) introduced it. The second important element was "engagement learning" and ten students (52.6%) mentioned it. The third important element included "learning environment" and "teaching techniques", and each one got 9 students' (47.4%) proposition. 8 students (42.1%) considered "heuristics teaching" method as the effective method that should be used in classroom-based teaching. 6 students (31.6%) suggested that "interaction" could facilitate teaching and learning.

3.3 Students' Experiences of Online Learning

Students' experiences of online learning were examined at the beginning of the course first. Among these 20 respondents, 38.3% of them had no any experience of online learning, but 61.7% students had.

Eight students mentioned their experiences of online learning were formed in terms of the different courses, i.e. the *Theory of Television*, *Analog Electronic Technology, Programming, Photoshop, Flash, Director, ASP, and Television Facilities*. Some of their descriptions are quoted and analyzed as follows:

> ***Fragad:*** *Tutor gave us a networked-based course, "Television Facilities". Then, he asked us to learn it by ourselves. Then, tutor gave us some explanations in the classroom and asked us to submit our assignments. The evaluation method combined assignments with the final exam. However, it is just boring!*

Mr. Fragad described his experience of online learning. What he introduced learning approach was similar to the blended e-learning. However, it was hard to say that it was real blended. The relevant learning activities, the interaction between tutor and students, and the relationship between students and students should be elaborated and refined. Tutor would not only leave students an online courseware, and then did nothing to facilitate their learning. This student also commented that his feeling about this online course was boring. In the Chinese higher educational field, many tutors try to change their traditional teaching approach. However, many reasons result in their efforts were not successful, such as did not know the essential, and had no background of new teaching approach. Teachers should attend the relevant training activities to improve their abilities of the new teaching approach.

> ***Gigi:*** *I learned programming from the web. However, it is hard to read through the screen. I do not like it. Therefore, I cannot insist on the further study.*

Miss Gigi introduced her experience of online learning which was acquainted from the web, because she learned programming from there. There were plentiful online learning resources on the web, such as online course, e-books, e-papers, and online database. Miss Gigi did not like to read them on the screen because it caused her tired and uncomfortable. Her comments presented the native aspect of online learning.

(1) Students' experiences of online learning

Students' experiences of online learning were examined at the end of the course, and the results were utilized for comparing how their experiences changed. The qualitative data were mainly analyzed including students' understanding, comments, suggestions, or critiques of online learning.

> ***Gigi:*** *From this course, I already knew more about online learning. I realised students should have autonomy and explicit learning objectives. The problem for online learning is the adaptive learning ability. On the other hand, online learning is easy to lose learning purposes. The essential elements of online learning are learners' autonomy and explicit objectives.*

At the beginning of the course, Miss Gigi thought that the effectiveness of online learning was very generalized and easy to exhaust learners. At the end of the course, her attitude to online learning got a big change. She was not only having a positive attitude to online learning, but also forming her critical reflection on it. She also proposed that the essential elements of online learning included autonomy and explicit objectives.

> ***Miriam:*** *Online learning is quite good! Except learners study in the class, they also can learn through the Web. Online learning can facilitate classroom-based learning. We can share our understanding via the online learning environments. I thought the essential elements of online learning are students' autonomy and the attraction of the issues on the web.*

Miss Miriam had positive attitude to online learning at the end of the course. She thought that classroom-based learning and online learning were complementary each

other. She concerned students' autonomy and the attraction of the issues on the web. Her perspectives on online learning already went far beyond the superficial issues that were formed at the beginning of the course, such as tutor's guide, FTF communication, and learning resources.

(2) The Effective Tools of the Online Learning Environments

In the second survey of the study, students introduced their thinking about the effective tools of the online learning environment. The data about the effective tools of the online learning environments is presented in Table 3.

Table 3. The effective tools in the online learning environments

Items	Number of students choosing		Items	Number of students choosing	
	Freq.	Per.		Freq.	Per.
Discussion forum	11	84.6%	Learning resources	1	7.8%
Chatting	5	34.5%	Blogs	1	7.8%
Email	7	53.9%	QQ	2	15.6%

"Discussion forum" got the significant response, which 84.6% students considered it as an effective tool for the online learning environment. However, at the beginning of the course, only 15% students chose this item, which illustrated that students liked to use discussion forum to express or share their understandings. Students also introduced their considerations to them.

> ***Kinston:*** *Discussion forum is an asynchronous communication tool. It can be used effectively for those which need long time to ponder.*

Mr. Kinston stated that discussion forum was an asynchronous communication tool. It could facilitate students' pondering if they took time for it.

> ***Onjcn:*** *Discussion forum can reflect students' reality in time and more detailed. The specialised discussion forum (sub-forum) results in the issues of discussion more concentrated and professional.*

Mr. Onjcn thought discussion forum could reflect students' reality in time and more detailed. He also mentioned the issues in discussion forum were more concentrated and professional.

The second significant item was "email" (53.9%). At the beginning of the course, only 25% students chose this item. This also demonstrated that students considered email was an effective tool when they were engaging in the online learning activities. 34.5% students chose item "chatting" at the end of the course. However, 40% students chose it at the beginning of the course. Some students changed their attitudes to the "chatting" when they acquired the new experiences of online learning. It revealed that "chatting" had some limitations, such as synchronous and the chatting contents could not be easily saved. 15.6% students thought that "QQ (Quick Quest)" was the effective

tool for online learning. Only 7.8% students considered "learning resources" and "blogs" were the effective tools.

4 Conclusions

This is a small-scale research on how Chinese students' perception of blended e-learning. It is only a short-term since blended e-learning has been using in the field of Chinese higher education, which is important to know how Chinese students' thinking and understanding of blended e-learning, especially for knowing how they experience this brand-new approach. The results presented in this study demonstrate that the students have diverse understanding and experiences of blended e-learning and their attitudes are positive. The practical experiences would be crucial for forming and improving students' understanding of blended e-learning.

Acknowledgements

I thank for the students who enrolled in this course. Without their cooperation, this study will not be finished. I also thank my PhD supervisor, Prof. McConnell for his direction and encouragement.

References

1. Smith, T.: Asynchronous discussions: Importance, design, facilitation, and evaluation (2001), http://www.ion.illinois.edu/pointers/2002_11/page1.html
2. Mason, R.: Guest editorial: blended learning. Education, Communication & Information 5(3), 217–220 (2005)
3. Stubbs, M., Martin, I., Endlar, L.: The structuration of blended learning: putting holistic design principles into practice. British Journal of Educational Technology 37(2), 163–175 (2006)
4. Ward, J., LaBranche, G.A.: Blended learning: The convergence of e-learning and meetings. Franchising World 35(4), 22 (2003)
5. Bersin & Associates. Blended Learning: What Works? E-LearningGuru (2003), http://www.e-learningguru.com/wpapers/blended_bersin.doc
6. Bielawski, L., Metcalf, D.: Blended eLearning: Integrating Knowledge, Performance, Support, and Online Learning. HRD Press, Inc., Amherst (2003)
7. He, K.: Examining the new development of educational technology based on blended learning (part I). Journal of Audio-Visual Education Research 131(3), 1–6 (2004)
8. Li, K., Zhao, J.: Blended learning: theory and application approach. Journal of Audio-Visual Education Research 135(7), 1–6 (2004)
9. Ginns, P., Ellis, R.: Quality in blended learning: Exploring the relationship between online and face-to-face teaching and learning. Internet and Higher Education 10(1), 53–64 (2007)
10. Delialioglu, O., Yildirim, Z.: Students' perceptions on effective dimensions of inter-active learning in a blended learning environment. Educational Technology & Society 10(2), 133–146 (2007)

11. Reichlmayr, T.: Enhancing the student project team experience with blended learning techniques. In: The Proceedings of 35th ASEE/IEEE Frontier in Education Conference, T4F-6, Indianapolis, IN, October 19-22 (2005)
12. Tian, R.: Exploring on organising method of English learning group. Journal of Sheyang College of Education 5(1), 53–56 (2003)
13. Gao, R., Kong, W.: Problem-based learning in the networked learning environment. Journal of Chinese Educational Technology 211(8), 28–33 (2004)
14. He, K.: Constructivism: instructional mode, method, and design. Journal of Beijing Normal University (Social Science) 143(5), 74–82 (1997)

Research and Application of Learning Activity Management System in College and University E-Learning

Li Yan[1], Jiumin Yang[2], Zongkai Yang[1], Sanya Liu[1], and Lei Huang[2]

[1] Engineering Research Center for Education Information Technology, Huazhong Normal University, Wuhan, China 430079
yanli@mail.ccnu.edu.cn
[2] Department of Information Technology, Huazhong Normal University, Wuhan, China 430079
yjm@mail.ccnu.edu.cn, zkyang@mail.ccnu.edu.cn, lsy5918@mail.ccnu.edu.cn, ccnuet@163.com

Abstract. LAMS (learning activity management system) is a very flexible learning design tool. With the combination of LAMS and the content, teachers/lecturers can easily organize the learning activities for the students in e-learning. This paper focuses on the application and functions of LAMS in college and university e-learning. Through formal research and evaluation of the impact of LAMS on e-learning in 3 disciplines (carried out in Huazhong Normal University), the presentation summarizes both the merits and the downsides of the application of LAMS in college & university e-learning and concludes with reflection on how further work will go on.

Keywords: LAMS (learning activity management system), College & University teaching, E-learning; Application research.

1 LAMS Introduction

Learning activity management system (LAMS) is a system for creating and managing sequences of Learning Activities. It was developed by Macquarie university, Sydney Australia. LAMS has four main areas: authoring, monitoring, learner and administration [1].

LAMS can support many activities like chat, forum, notice board, notebook, Q & A, share resources, submit files, survey, voting etc. Some activity can be finished by learners themselves and some need to be finished by the learners' collaboration. Teacher can use drag-and-pull method to assemble activities into sequence easily. It is a good tool for designing, managing and delivering online collaborative learning activities. It provides teachers with a highly intuitive visual authoring environment for creating sequences of learning activities [2]. LAMS can be used as a stand alone system or in combination with other learning management system such as Moodle, WebCT or BlackBoard.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 171–179, 2008.

2 Theory Foundation of This Research

Although the concept of learning design appears very frequent recently with the rapid development of e-learning, it is, in fact, far from a new idea. Whether people realize it or not, learning design occurs whenever and wherever teaching or learning happens. The instructors design for learning consciously or reflectively in teaching-related activities and the learners make learning design decisions consciously or subconsciously in learning-related activities.

Learning design can be conceptualized at three levels – theory, standards and software. At the highest level, Learning Design theory is based on the general idea of people doing activities with resources/environments (e.g., Sloep, 2002) [3].At the second level, there are as yet no formally ratified technical standards for Learning Design. And at the third level, there are a number of software systems in use or in development that are based on Learning Design theory [4].

The process of learning design involves the definition of learning objectives, the development of the narrative description of learning and teaching scenario, the creation of the learning activity workflow from narrative description, the assignment of resources, tools and people to activities, the running(real-time),the learner support and on-the-fly adaptation, and reflection(including sharing outputs for peer reflection) [3].

Apparently, such a complex process demonstrates very well the need for software tools to be developed to facilitate the instructors and the learners. And among all the currently available software tools, LAMS is one of the best and most popular learning design supporting tools [5]. In fact, the initial feedbacks from LAMS users have proved it to be an effective tool since it provides a new approach to sharing and adaptation in e-learning.

3 The Application Scheme Design and Implementation about LAMS in College and University E-Learning

3.1 Research Objective

In order to research the application about LAMS in higher education e-learning, we chose three representational courses in Huazhong Normal University in 2007. They are “Modern Educational Technology”, “Advanced Multimedia Technology” and “New Horizon College English, Book II”. In these three courses, we used LAMS as learning platform with some learning contents. Compared with traditional learning methods in classroom, this research aims to discover and obtain the application model and functions of new e-learning method with LAMS through our research. Because the Chinese version of LAMS is not completed, we use English version of LAMS Version 2.0.4.

3.2 Research Scheme

Table1 gives us the scheme of our research.

Table 1. Research scheme

Curriculum name	Modern Educational Technology	Advanced Multimedia Technology	New Horizon College English (Book II)
Curriculum character	public required course of all the students with normal major	professional optional course of information management department	public required course of all the students in the university
Students' grade	Grade 2004, Grade 2005	Grade 2004	Grade 2006 from Chinese department
Students' major	all kind of majors	information management and information system, electronic commerce	Chinese Language and Literature
Learning content with LAMS	e-learning resources, environment and methods	chapter3, the capture and transaction of video signal	unit 3, unit 4 and unit 5
Students	class 2, arrange by random, plan 25 students, actually 14 students finished class 4, sign up by volunteer, sign up 29 students, actually 20 students finished	whole class, total 30 students, actually 22 students finished	whole class, total 35 students
Time arrangement	one month (if this content is learned in classroom, the plan time is 4 learning hours)	two weeks (if this content is learned in classroom, the plan time is 3 learning hours)	
Apply method	network-based completely, learners with LAMS needn't go to classroom with this content.	network-based completely, learners needn't go to classroom with this content.	Use LAMS as the aid of traditional learning, students use LAMS before, in and after class according to the teacher's arrangement.
Data collection	with the "survey" activity in LAMS, design questionnaire. Take back valid questionnaire 34 pieces.	with the "survey" and "voting" activities in LAMS, design questionnaire. Take back valid questionnaire 22 pieces.	The teacher collected data, advices and suggestions through the communion with students.
Learning activities	totally there are 26 activities in one sequence including 11 activities with interaction (chat, forum, Q & A, share resources, survey etc.). see Fig.1.	totally there are 27 activities in one sequence including 12 activities with interaction (chat, forum, Q & A, share resources, survey etc.). see Fig.2.	12 activities in unit 3 sequence, 5 activities in unit 4 sequence and 5 activities in unit 5 sequence including some activities with interaction. The learning sequence in unit 5, see Fig.3.

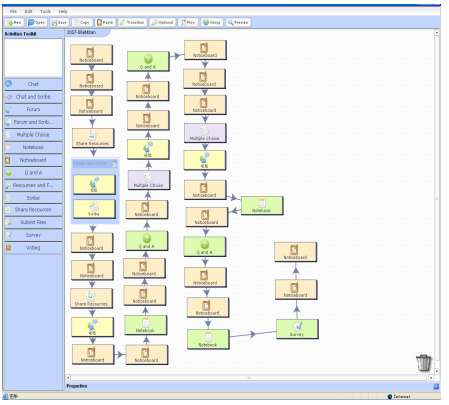


Fig. 1. Learning sequence in "Modern Education Technology" course

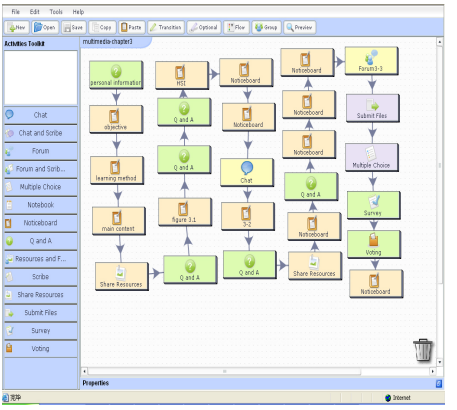


Fig. 2. Learning sequence in "Advanced Multimedia Technology" course

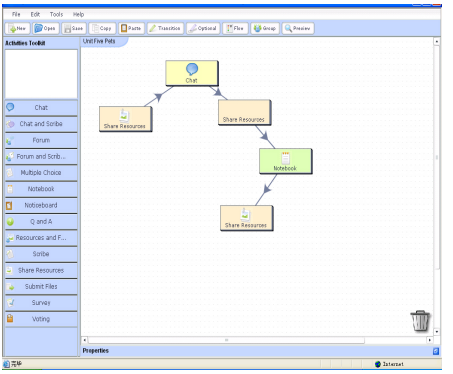


Fig. 3. Learning sequence in "New Horizon College English, Book II" course, unit 5

4 Data Analysis

4.1 Data Analysis in "Modern Educational Technology" Course

4.1.1 Motivation

Question: Why did you sign in the activity about LAMS experience? (multiple choice)

Result: see Table 2.

Table 2. Result about motivation

Nomination:	Total votes
Good chance to experience new online learning method	16
Sounds fresh, just try it	15
Don't need to go to classroom	4
It is my right to sign in attending new experience	1
other	3

From the result, we can see the students' learning attitude and motivation are active and correct. The new learning method can inspire students' interesting.

4.1.2 Difficult Degree

In order to get the result of influencing factors when learners use LAMS with quantity value, we set: 0 stands for "not difficult"; 1 stands for "a little difficult"; 2 stands for "very difficult". The statistic result sees Table 3.

Table 3. Statistic data about influencing factors' difficult degree

index	Nomination:	Statistic value
1	Bad connection with LAMS or other websites	1.26
2	Can't effectively evaluate the learning result	1.00
3	Feel lonely when study	0.91
4	Not convenient for using computer	0.85
5	Not adapt learning online	0.74
6	Not good at managing learning time	0.62
7	Not familiar with LAMS platform	0.38
8	Not good at English	0.26

From the result, we can see, the network infrastructure is the biggest problem when learners use LAMS. It could be solved by the support of some departments of university like computer center. For the factor 2, it needs the platform developers to enhance the evaluation function of LAMS. For the factor 3, 5, 6, they are common problems towards network learning. We need to enhance students' ability and literature about online learning. And we can see, for most college students, using English version LAMS is not the problem.

4.1.3 Time Spend

Question: Compared with traditional learning method in classroom, the time you spend on learning the LAMS sequence is _____.

We set: 2 stands for "much longer"; 1 stands for "longer"; 0 stands for "almost same"; -1 stands for "shorter"; -2 stands for "much shorter"

The statistic data value is: 1.21

From the result, we can see the learning time with LAMS is much longer than learning in classroom for most students. But it is just the time factor, not the effect result.

4.1.4 Effect

The data about application effect with LAMS sees Table 4.

Table 4. Statistic data about application effect with LAMS

	Condition set:	**Statistic value**
(1) Compared with learning in classroom, your thinking is:	2 stands for "much more active"; 1 stands for "active"; 0 stands for "almost same"; -1 stands for "passive"; -2 stands for "much more passive"	0.62
(2) Is this online learning with LAMS helpful for you to understand Educational Technology?	2 stands for "very helpful"; 1 stands for "helpful"; 0 stands for "not helpful"	1.29

The result illustrates that for most students, learning with LAMS can make their thinking more active and it is helpful for them mastering the knowledge and understanding the concepts better in their professional field. In one word, with LAMS this new e-learning platform, students can get better learning effect than traditional learning method.

4.1.5 Attitude

Towards learner's attitude, we set two questions let students voting. The first question is: Which learning method do you prefer? The second is: Will you volunteer to use LAMS later? The Fig.4 and Fig.5 can give us intuitionistic expression about the result.

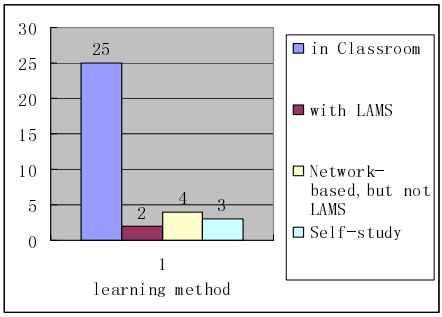


Fig. 4. Learning methods the students prefer ("Educational Technology" course)

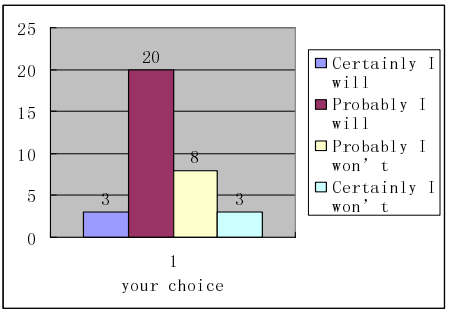


Fig. 5. The students' choice towards LAMS ("Educational Technology" course)

The results sound conflict. Although most learners think using LAMS for e-learning is better than traditional learning, and they probably use this method again, they would like to go to classroom and learn knowledge from teacher face to face. We analyze the reasons come from these factors: one, the learning sequence is too long. Students should spend much time on it. This would cause that students feel tired. Two, the students are not very familiar with LAMS platform, and without the teacher's direct guide and help, students will feel helpless sometime. Three, every

new thing needs time to be adapted by people. For this new learning method, students need some time to adapt it and like it.

4.2 Data Analysis in "Advanced Multimedia Technology" Course

Through "survey" and "voting" tools, we collected students' feedback and got the statistic data about LAMS application result in this course. See Table 5, Fig.6. and Fig.7.

Table 5. Some statistic data with LAMS application in this course

	Nomination	total votes
(1) Time spend	Less than 3 learning hours	7
	3 to 5 learning hours	12
	More than 5 hours	3
(2) With LAMS, how do you think about the learning effect?	With LAMS, the effect is better than traditional learning.	12
	With LAMS, the effect is less than traditional learning.	2
	No obvious difference.	6
	Open response	2

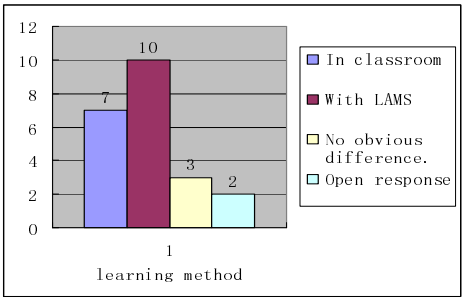


Fig. 6. Learning methods the students prefer ("Advanced multimedia technology" course)

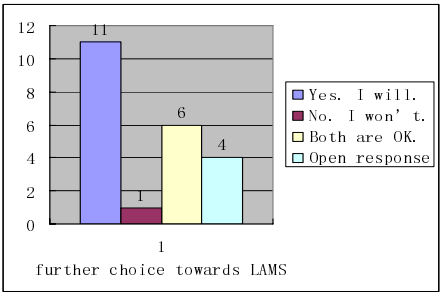


Fig. 7. The students' choice towards LAMS ("Advanced multimedia technology" course)

From the feedback of students, we can see: using LAMS as e-learning platform, students should spend much time than using traditional learning method. But, students could receive more resources and obtain more knowledge towards same knowledge concept. We don't deny that with LAMS platform, there are some problems like the interaction between teacher and students is affected by the limitation of time; when students met problems, they could not get rapid response from teacher sometimes; there are no audio and video instant interaction with LAMS; it is not convenient for teacher demonstrates some technical operations with LAMS to students. However, most students think that LAMS is a good e-learning platform and they can get better learning effect with LAMS. The best advantages of LAMS from students' opinion are

represented with flexible learning time, self-learning method, colorful and rich resources, good interaction etc.

4.3 Data Analysis in "New Horizon College English, Book II" Course

In this course, teacher got students' feedback through communion with them.

Advantages in the eyes of students include: (1) New teaching and studying mode arouses students' interests; (2) The application of LAMS is close to the life of modern students; (3) New teaching concept enriches the varieties of classroom activities; (4)New classroom activities can explore students' potentials, get students involved in more activities leading to the use of the language; (5) LAMS can strengthen the interaction between the teacher and students, thus a better learning environment is established.

Disadvantages of e-learning with LAMS include: (1) The system is web-based, however, the speed of the internet and other hardware can not meet the requirement; (2) The system is not stable and students can't make use of the system at the same time; (3)The account numbers of students are hard to manage; (4) The teacher is not familiar with all the skills in designing the sequences; (5)This course focuses more on spoken English, while the system only can offer the communication based on written English.

5 Evaluation and Future Work

The outcomes differentiate when LAMS is implemented in different discipline or classroom. Except the effect of LAMS, they are changed with the teachers' different perspective on the design of the learning content and the implementation of teaching.

Overall, the feedback from LAMS users has been very positive. It adds value to the teaching and learning practice in e-learning by creating a friendlier and more relaxing learning environment, fostering a better understanding and interaction between teachers and students, and improving the efficiency and flexibility of the teaching and learning activities [6].

Yet, whilst all the upsides outlined above, there are problems in the application of LAMS. The implementation is limited since it is an online web-based system that runs through a standard browser capable of supporting Flash and it is still quite limited in its functionality and flexibility.

Thus, further work should be conducted. We should guarantee the availability of the network, provide the platforms for e-learning, and offer necessary training and instruction for the teachers. And since the real application of LAMS in e-learning is limited and it still appears most in some educational researches, it is recommended that ongoing work must be conducted to discuss how to apply LAMS effectively to the real practical learning process.

Acknowledgement

This research is supported by the Program of Introducing Talents of Discipline to Universities，Ministry of Education and State Administration of Foreign Experts

Affairs of China（NOB07042）, and National Great Project of Scientific and Technical Supporting Programs (NO. 2006BAH02A24).

References

1. Dalziel, J.: Implementing Learning Design: The Learning Activity Management System (LAMS). LAMS™ Teacher's Guide, November 2006, V2.0. Copyright © 2004-06 James Dalziel, pp. 75–80 (2006)
2. About LAMS, http://www.lamsinternational.com
3. Dalziel, J.R.: From re-usable e-learning content to re-usable learning designs: Lessons from LAMS [EB/OL] (2004)
4. http://www.lamsinternational.com/CD/html/resources/whitepapers/Dalziel.LAMS.doc
5. Britain, S.: A Review of Learning Design: Concept, Specifications and Tools[EB/OL] (2004), `http://www.jisc.ac.uk/uploaded_documents/ACF1ABB.doc`
6. Russell, T., Varga-Atkins, T., Roberts, D.: Learning Activity Management System (LAMS) Specialist Schools Trust Pilot Review. CRIPSAT, Centre for Lifelong Learning, University of Liverpool. BECTA ICT Research (2005)
7. Marshall, S.: Leading and managing the development of e-learning environments: An issue of comfort or discomfort? In: Atkinson, R., McBeath, C., Jonas-Dwyer, C., Phillips, R. (eds.) Beyond the comfort zone: ASCILITE 2004. Proceedings of the 21st Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education, Perth (2004)
8. Kenny: A research-based model for managing strategic educational change and innovation projects. In: Annual Conference Proceedings of HERDSA (the Higher Education Research and Development Society of Australasia), University of Canterbury, Christchurch, NZ, 6-9 July (2003)
9. Huang, R., Zhou, Y.: The Characteristic Analysis of distance learning. China Audiovisual Education, pp. 75–79 (March 2003), pp. 69–71 (April 2003)
10. Wang, Y., Peng, H., Huang, R.: Scale of Distance Learners' Learning Motivation Development and Initial Application. Open Education Research, pp. 74–78 (May 2006)

Motivate the Learners to Practice English through Playing with Chatbot CSIEC

Jiyou Jia and Weichao Chen

Department of Educational Technology, School of Education,
Peking University. Beijing, 100871, China
jjy@pku.edu.cn, st510@gse.pku.edu.cn

Abstract. CSIEC (Computer Simulation in Educational Communication), is an interactive web-based human-computer dialogue system with natural language for English instruction. In this paper we present its newest developments and applications in English education. After brief introduction of the project motivation and the related works, we illustrate the system structure with a flow diagram, and describe its pedagogical functions in details, including free chatting, chatting on a given topic and the chatting scoring mechanism. We review the free Internet usage within six months, and evaluate its integration into English classroom. The summarization and assessment findings confirm that the chatting function has been enhanced and fully used by the users, and the application of the CSIEC system in English instruction can interest the learners to study English and motivate them to practice English more frequently. Finally we discuss the application driven approach of system development, and draw some conclusions for the further improvements.

Keywords: CSIEC, English Learning, Chatting, Playing, Scoring, Motivation.

1 Introduction

1.1 Motivation

English, as an international language, is treated as a key tool for the development and cultivation of the cross-cultural communication ability. English language is now listed as one of the three core courses in China's elementary and secondary education, and as a compulsory course in higher education. Statistical data shows that there were more than 176 million people learning English in China in 2005 [1].

However, some problems exist in the English education in China. First of all, one of the best ways to learn a foreign language is through spoken dialogue with native speakers. But it isn't practical in the classroom due to the one-to-one student/teacher ratio it implies, especially in China and other countries with English as a foreign language. A number of factors ranging from the lack of time to shyness or limited opportunity for quality feedback hamper using the target language [2]. The language environment and few qualified English teachers in China can't supply enough chance of authentic talking. So school teachers often complain of working burdens, and don't

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 180–191, 2008.

have enough time to converse with students in English. Secondly, although learning English through communication and application has been emphasized recently, passing examinations is the main motivation of many students to learn English. Thirdly the grammar instruction is crucial to China's English education, because Chinese differs greatly from English in grammar [3]. Without basic grammar knowledge, the students can't make great progress, as they mostly only practice English in school time, and can't learn it spontaneously from the social environment.

A potential solution to these problems is to apply computer spoken dialogue systems to role play a conversational partner. If we could design an interactive web-based system which could chat with the English Learners anytime anywhere, their great demand for learning partners could be fulfilled. Such a system should aim at helping the learners improve their skills of using English through frequent chatting with them in English, as well as encouraging them through playing and scoring mechanism. Motivated by the great demand for English instruction, we in 2002 began to design such a system. Our design principle is application and evaluation oriented. So long as the system is applicable, we put it into free use in the Internet and get the user feedback. We also cooperate with the English teachers and integrate the system into English instruction. Through the systematic application and evaluation we get more suggestions and critiques, which can direct our research more effectively.

1.2 Related Works

Brennan defined a chatbot as "an artificial construct that is designed to converse with human beings using natural language as input and output" [4]. ELIZA [5], the first chatbot, used key words to analyse input sentence and created responses based on reassembly rules associated with a decomposition of the input. The syntactic way of NLP (Natural Language Processing) exemplified by ELIZA has been developed significantly from 1960s up to now, leading to the development of various chatbots. Since 1990s with the improvement of natural language processing, chatbots have become more practical, and have also been applied in education.

Graesser et al. [6] used "AutoTutor", an intelligent tutoring system with mixed-initiative dialogue simulating a human tutor via conversation with the learner in natural language, to enhance the learner's engagement and the learning depth.

Seneff [7] described several multilingual dialogue systems designed to address the need for language learning and teaching. A student's conversational interaction was assisted by a software agent functioning as a tutor with translation assistance anytime.

Kerfoot et al. [8] described an experimental use of chatbots as a teaching adjuvant in training medical students. Their web-based teaching using chatbots increased test scores in four topics significantly and learning efficiency three-folds.

Abu Shawar and Atwell [9] developed algorithms for adapting a chatbot to chat in the language and topic of the training corpus. The evaluation feedback from language learners and teachers indicated that these adaptive chatbots offered a useful autonomous alternative to traditional classroom-based conversation practice.

Kerly et al. [10] described an experiment to investigate the feasibility of using a chatbot to support negotiation. Its result showed that most students liked the chatbot as the chatbot helped them understand their learner model.

The related works above show the usage of chatbot systems in education is drawing more attentions from related researchers. This trend confirms our determination to further the development of the CSIEC system and its application in English education.

2 System Compositions and Technologies

In the system design, contrary to the partial parsing adopted in many other systems, we attempt the fully syntactical and semantic analysis of the user inputs, as the logician G. Frege pointed out: "The meaning of a sentence exists in the meanings of all words within the sentence and their conjunction method" [11]. After parsing the user input we obtain the user information in the form of XML, i.e. NLML and call them the user facts. The facts are retrieved from natural language expressions, and also represented with the annotation of natural language in the sentence ontology. These facts function as the main contextual source of the robot dialogue reasoning. This thought originates from L. Wittgenstein's theory (1918/21) about the world, facts, objects and human language: "The world consists of facts, the facts consist of objects. The facts are reflected in the language. A logical picture of facts is a thought." [12]

The current CSIEC system is version 9. The whole system is mainly made up of the following components, which are illustrated in Fig. 1.

A. HTTP request parser resolves the user request from http connection and gets some parameter values: input text, scenario topic, agent character, speech speed, spelling and grammar checker, etc.

B. English parser parses the user text into NLML (Natural Language Markup Language). NLML is a dependency tree in XML form, and structurally labels the grammar elements (phrases), their relations and other linguistic information.

C. NLML parser parses the NLML of the user input into NLOMJ (Natural Language Object Model in Java) which represents the grammatical elements and their dependency with the Sentence ontology in the working memory [13]. Through NLOMJ the declarative sentence is retrieved and decomposed into atomic facts consisting of only one subject and one verb phrase.

D. NLDB (Natural Language Database) stores the historical discourse, the user atomic facts in the form NLML, the robot atomic facts which are also expressed in NLML, and other data.

E. World model contains common sense knowledge which is the basis for response generation and logical inference. It is now represented by WordNet [14].

F. CR (Communicational Response) mechanism comprehensively takes into accounts the user facts stored in NLDB, the world model, the personality of the user expressed in the previous dialogue, and that of the robot itself selected by the user.

G. Scenario dialogue handler creates the robot output corresponding to the user input within a given scenario.

H. Scenario show handler creates the random robot-robot talk show scripts within a given scenario.

I. Scenario DB stores the robot-robot talk show scripts and human-robot dialogue scripts which are manually written by designer, for example English language teacher.

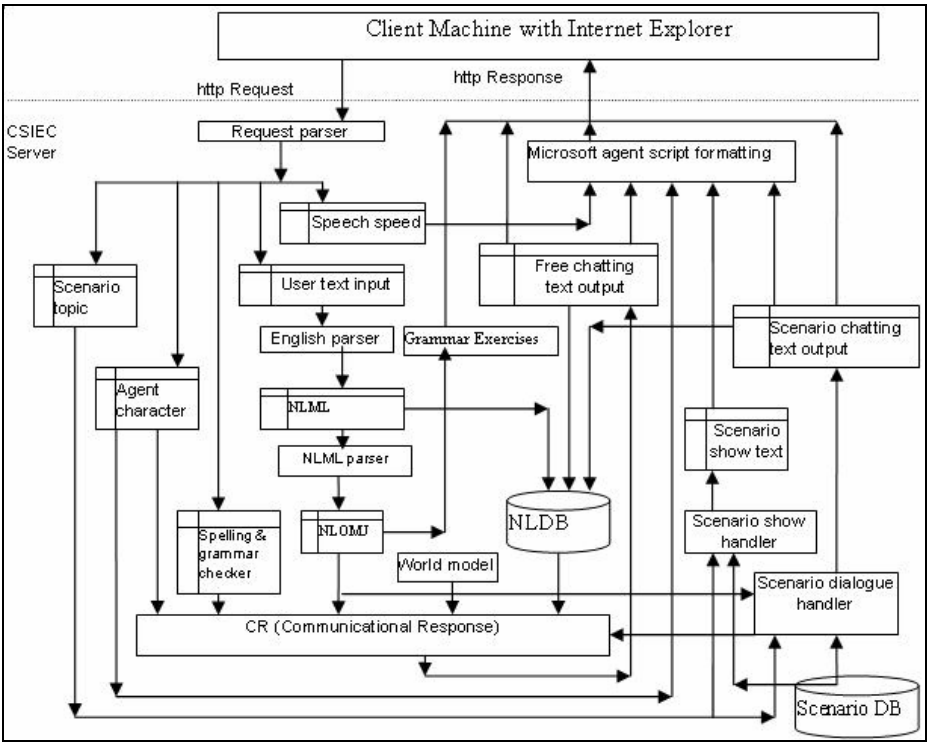


Fig. 1. The compositions of CSIEC system

J. Microsoft agent script formatter transforms the output text into VB scripts, considering the selected agent character and speaking speed.

K. Browser/Server interface processes the http request from client machine and responds with the robot output, either in text or with VB script.

3 Functions: Chatting, Playing and Scoring

3.1 Multimodal User Interface and Selectable Chatting Pattern

Human-computer dialogue in natural language is the most specific function of the CSIEC system. As in human being's authentic dialogue situation, the Internet users have various preferences for the dialogue simulation. In order to adapt to variant user preferences the CSIEC provides several user interfaces and dialogue patterns.

First of all the user can chat with the robot either through text or via voice. The users can hear synthesized voice and watch the avatar performance through Microsoft agent technology. They can also speak to the robot through a microphone with the help of a third party program like IBM ViaVoice.

Secondly the robot can check the spelling and grammar of the input text upon the user's request.

Thirdly the chat topic between the user and the robot can be either free (unlimited) or specific (limited). The unlimited dialogue simulation doesn't specify the dialogue topic and content. It suits the need of the users whose English is fluent or who are at least good at written English, as well as users who are extroversive or conversational. However, users whose English is poor, or who are introversive, have little to chat with the virtual chatting partner. For those users, an instructive dialogue in a specific scenario guided by the agent is more helpful.

In normal human talking these two chatting patterns are not absolutely separated, but often interleave each other. Our system considers this interaction too. In the next 2 subsections we introduce the two patterns in details, as well as their relationship.

3.2 Free Chatting Adaptive to User Preference and Topic

In the free chatting the users with different characters and personalities may choose different types of chatting pattern. For example, some users may prefer to chat with someone who listens to them quietly most of the time; while some others may hope the chatting partner can tell stories, jokes or news. For the sake of user dialogue personalization we designed five Microsoft agent characters which represent different kinds of chatting patterns [15]. Christine always tells the user stories, jokes and world news. Stephan prefers to listen quietly when the users share with him their own experiences. Emina is a curious girl, and is fond of asking users all kinds of questions related with the users' input. Christopher provides comments, suggestions and advices on the user's input. Ingrid behaves as a comprehensive virtual chatting partner, who gives users responses corresponding to both the input text and the discourse context.

Upon user registration to the chatting system the user's profile is obtained and recorded, such as the gender, birthday, educational level, and province. So the corresponding chatting topic and content can be generated based on the personal information. Certainly if the user wishes to change the chatting topic during the process of the robot's narrating comments or asking questions, the robot should terminate this process and transfer to another topic. If the user specifies a topic, for example, "*I want to talk about sport*", the robot changes the topic to it. If the user just expresses the wish to change the topic, but doesn't determine a topic, such as "*I want to talk about another topic*", the robot selects one from the waiting topics list which has not been talked about with the given user.

The user's interests are also expressed in the input texts, e.g. the mentioned nouns and verbs in the sentences. So the chatting topic can be triggered by nouns and verbs, and their combination. More frequently one noun is or several related nouns are talked about, the related topic is more emphasized. The chatting between the user and robot can be regarded as guided chatting or chatting in some context.

Then we deal with the chatting on a given topic in two ways. The first one is predefining some comments or asking some questions about this topic. By talking about this topic only one statement or question will be randomly selected and given out. The second way is to search the topic or related topic in the guided chatting within a given scenario, and then transfer the chatting to the guided chatting in a given scenario, which will be introduced more in next subsection. In Figure 1 the arrow from the scenario dialogue handler to the communicational response indicates this relationship.

Summarily the goal of free chatting is to motivate the user's talking desire. For this purpose the robot tries to adapt itself to the user's interest, and launch new topics.

3.3 Guided Chatting in a Given Scenario

We then explain our approach to guide the chatting with the robot in a given scenario. The dialogue should be developed step by step around a red line or a topic for this scenario. Due to the extreme complexity of natural language this dialogue development is exceedingly nonlinearly complicated. It can be described by a complex tree structure with many branches. These branches can be pragmatically and semantically countable, but syntactically uncountable. So this dialogue scenario tree is much more complicated than the classical chess decision making programs, which has finite state change selections in every step.

We use scripts to describe the decision tree in the dialogue on a given topic. The script is made up of lines of dialogue steps (states), every of which is a branch in the decision tree. Suppose the robot speaks at first. In every line there must be the text output from the robot and its order number in the dialogue. This output may be triggered by specific user input, which we call the prerequisite of this output text. The robot may also expect the user inputs certain texts, or some texts with specific semantic or syntactical characters, which we call the expectation of this output text. We write the line in the script with the following format:

Nr. <prerequisite> (text) <expectation>

The "Nr." and (text) are the necessary two components in every line. The "Nr." is an integer which indicates the line order in the whole script, whereas the "text" can be any text from the robot, either statement, or question, and so on, and it is written within closed brackets.

In a script line the prerequisite and expectation are optional. If they appear they must be written within closed sharp brackets. If the prerequisite exists and is satisfied the output text can be given out by the robot. The expectation means the robot hopes the user responses to this text with some specific syntactic and/or semantic features, and can be applied to instructional goal. For example if the user's input does not satisfy the robot's expectation he/she will face the previous robot output again, until the expectation is fulfilled. This dialogue pattern can be used for drill. Another alternative is that the user is given a high mark if his input satisfies the robot output, otherwise a low mark, although the robot continues the next dialogue. This pattern can be used in test or examination.

The format of the prerequisite is:

<Nr, variable 1: value 1, value 2...; variable 2: value 1, value 2...>

The format of the expectation is:

<variable 1: value 1, value 2...; variable 2: value 1, value 2...>

Both are almost the same form. Only the prerequisite needs an order number indicating the expectation of which line this condition fulfills. There may be more than one value for a given variable. This means if the variable equals any one of the listed values, the condition is fulfilled, i.e. the values for a given variable have the relation of logical disjunction. There may be also more than one variable and its corresponding values. The relation among these variables is logical conjunction.

This discourse script is difficult to be written by normal authors, for example the English teachers who want to use this program to train the students. Even an error with a bracket will cause the misunderstanding of the computer program. Thus we have designed a Java GUI, i.e. DSE (Discourse Script Editor) for editing the scripts step by step more easily. With it a normal user such as an English teacher needs not to pay attention to the writing format, but to the discourse content and process. However, he/she has to spend much time on planning this discourse script between the robot and the human user, just like a film director. This work is not just the language teaching, but also the teaching of response strategies through natural language.

3.4 Listening Training

We use the Microsoft agent technology to synthesize the output text, because the agent's voice is lifelike, the agent's figures, movements as well as actions can be designed very vividly, and it can also synchronously display the spoken text, which facilitates the aural understanding and activates the user's interests. We have also designed seven facial expressions (neutral, happy, sad, feared, disgusted, angry and surprised) for every agent character and hope the textual emotional expressions can be accompanied by the agents' facial changes. The robot's reading speed can be adjusted by the users at any time. We have also designed a text-reading webpage where the agent can read any texts inputted or pasted by the user.

Different from the traditional audio technologies such as audio players, the user confronts with unexpected robot text and voices, just like talking with a real human being. So it is hoped that this function can benefit the user's listening comprehension and prompt response.

3.5 Talk Show of Two Robots

This function is designed to aid the user's chatting on a given topic. With it the users can watch the talk show of two robots before the human-computer interaction. The talking texts are predefined by the teacher for the specific context or topic. However, the actual texts for a given meaning can be expressed randomly. So this kind of talk show is different from the monotone one presented in the traditional video or audio cassette. It will enforce the learner's spontaneous listening and understanding. The talk show script texts can be readily written by the teachers with any text editor.

3.6 Automatic Scoring of Gap-Filling Exercises without Well-Defined Answers

Traditional computer-based gap filling exercises require a definite answer or a set of definite answers. For the questions whose answers are difficultly to be listed, the human manual check is still unavoidable. However, this kind of exercise without predefined answers can advance the creative thinking of the students.

With the spelling and grammatical check function the CSIEC system can decide if a filled gap-filling sentence is grammatically correct. Therefore it can be applied to automatically assess the gap-filling exercises and relieve the teachers' burden. So currently the system provides the interface for teachers to design new gap-filling

exercises, as well as the interface for learners to do these exercises and then to get the automatic assessment results.

An example of gap-filling exercises is: "*I ( ) a student*." The correct answer to the gap can be: "*am", "want to be", "will be", "have been", "need", "help"*, etc.

3.7 Scoring Mechanism

In order to motivate the users to learn English we trace users' usage of different functions and give them certain scores. The scoring principle is encouraging the usage of chatting with agents, and with spelling and grammar checking. By the chatting on a given context, the user is given a high mark if the input satisfies the robot output, otherwise a low mark. This mark also contributes to the total score.

The user can review his performance and scores after entering the system. This function is very important and helpful for self learning. A special user who is labeled as the teacher can access the performance and scores of all the users who are classified as his/her students. This automatic monitoring function is very necessary for the teacher to assess the students' learning behavior and progress.

4 Application and Evaluation

4.1 Summative Evaluation of Free Using in Internet

The internet users get to the CSIEC website (www.csiec.com) mainly through search engines, because our website has become one of the top 5s in the searching results of famous search engines such as google.com, yahoo.com and baidu.com by related keywords such as "chatbot", "English chatbot", "Online English learning" in Chinese or in English, although we haven't made any large-scale advertisement. The effectiveness and attractiveness of the system's adaption to English learning in China has been somewhat demonstrated by this practical achievement.

With the human-computer dialogues recorded in the database, we make a summarization of the system's chatting function from Jan. 20th 2007 to June 20th 2007. The different users who accessed the CSIEC during this period count 1783. The analysis of the demographic distribution of the users shows that more than half of the users are undergraduate students. The second large user population is middle school students. Except 45 students required to use the system in the evaluation, there are still 377 free users. Totally more than 80% of the users are different kinds of students.

4.1.1 Dialogue Duration

The chatting quality can be measured by the chatting duration between the user and the robot. To calculate the chatting duration we define two terms: round and number of the rounds. A round means a user input and a corresponding robot output to the user. Therefore the total rounds of a given user cover all dialogs between the user and the chatbot, and can be used to describe the duration of the user's chatting with this chatbot. We divide the number of the rounds into 4 classes, as Table 1 shows.

Table 1. The relation between the duration of dialogues and number of users

Dialogue duration	Range of the rounds numbers	Number of users	Number of users/ Total user number	Number of users/ Total user number in [16]
Short	(0, 10]	871	48.85%	62,34%
Long	(10, 50]	685	38.42%	30,10%
Longer	(50, 100]	136	7.63%	4,78%
Very long	(100, 580]	91	5.10%	2,79%
Total user number		1783	100.00%	100,00%

The average rounds number is 27.4 (48840/1783). The number of the rounds from each user varies from 1 to 580. From table 1 we draw the conclusion that c.a. 49% of the users chat with the robot briefly (<=10 rounds); c.a. 46% (38.42%+7.63%) of them chat with it long or longer; and only few, c.a. 5%, chat with it very long (>100 rounds). Compared with our previous finding in [16] which is listed in the last column of Table 1, the percentage of the brief chatting with the robot has decreased by 21.78%. Proportionally the percentage of the long and longer chatting has increased.

4.1.2 The Distribution of User Chatting Patterns

The CSIEC system provides multimodal user interface and selectable chatting patterns. Thus we investigate the distribution of chatting patterns. 84.7% of the chatting is held with the free chatting pattern, and only 15.3% uses the chatting in a given context. The reason may be that the free users do not understand what is chatting in a given context very well so that most users of the context chatting are the students in our project English classes. Among the chatting for a given context the text pattern is used almost as frequently as the agent pattern. It can be explained by our team's assistant and tutoring about the system usage, especially the installation and usage of Microsoft agent characters, in every unit of the English class.

Among the free chatting more users select the text version instead of agent version. One reason may be that the text pattern is more simple and convenient than the agent version, as the unskilled computer users may encounter some setting problems, what is proven by some users' feedback complaining that they can't use the agent version.

Among the free chatting the chatting without spelling and grammar check (c.a. 66%) is much more used than with check (c.a. 18%). This result reflects most free users treat the system as a chatting partner, so they'd like to chat with it more fluently instead of worrying about grammar and spelling errors. Human-computer chatting is the most unique function of the CSIEC system, therefore the users like to fully use it.

4.1.3 User Feedbacks

In the foot of almost every webpage of the CSIEC system we leave a feedback text area so that the users can straightforwardly enter their comments, critiques and suggestions. Through analysis of the user feedbacks we find as many critiques as praises. For example there are the following positive comments:

The robot is more advanced than before, and also personalized.
The access speed is faster than before.

The dialogue is fluent. I hope the master to enrich the robot's language.
The kind of communication can improve our English.

The negative comments point out either technical problems or content shortcoming. Some complain that they can't use the agent version or the agent voice sounds curious. Other problems include: the access speed is too low, the robot response is too slow, the dialogue for a given context is too short, in free chatting the robot always repeat a same sentence, etc. These problems should be tackled in the further improvement.

4.2 Formative Evaluation of English Class Integration

After discussing with the English teachers about the class integration and evaluation of the CSIEC system we came to a decision that the instructional instruments are on one side the talk show by two chatting robots and on the other side students' talking in English with one robot on a given topic corresponding to the textbook content. The main application goal is to facilitate role-playing activities in the English classes.

45 high school students in Grade 2 attended the study, and the teacher required the students to use the system together in the computer room. For the 10 units course content we designed 40 scenario scripts for the role-play talk show and human-robot chatting. During the whole term we formally evaluated it through questionnaires, observations in the classes, surveys with teachers and students focus groups.

The survey contained 6 items about the students' attitude toward the CSIEC's application in English instruction: enhancing fluency of English, enhancing confidence in communication, enhancing learning interest, mastering practical expression, improving listening skills, and reviewing key points in course units. All they were measured with a five-points Likert agreement scale, i.e. the value 5 indicates the maximum best agreement, and 1 is no agreement. The mean is 2.5, 2.8, 3.3, 3.2, 2.9, and 3.3, respectively. It shows that high school students feel the CSIEC-based English learning can help with course unit review, make them more confident, improve their listening ability, and enhance the interest in language learning.

Another item in the questionnaire shows 60.5% of the students "*liked*" or "*liked very much*" such a form of English learning, whereas only 2.3% disliked it. 60.5% of them would continue using the system after class, even without the teacher's request.

Through the integration and assessment of the system in English class instruction, some new functions have been added to the system according to the students and teachers' suggestions and comments. They include the adjustment of speaking speed of the agent character, two robots talk show, unlimited gap-filling exercises, etc. Thus the application and evaluation guide the development of the CSIEC system in the direction of users' practical learning needs.

5 Conclusion and Discussion

The original goal of the system is supplying a virtual chatting partner for English learner. So the chatting is the most fundamental function. The statistical analysis about the users' behavior indicates that the users have a preference for chatting without spelling and grammar checking. This fact proves that the users prefer the

unique chatting function which is lacked in other systems. So we must continue to reinforce this primary utility.

The chatting quality can be somewhat demonstrated through the chatting length. The increased percentage of the long and longer chatting shows that the free chatting quality of CSIEC is becoming better. The underlying design principle, i.e. fully syntactical and semantic analysis of the user input, and communicative response mechanism, as well as the effort of chatting personalization and adaptation contribute to that quality progress. Certainly the content analysis of the dialogues should also be conducted furthermore in order to investigate the chatting quality more exactly.

The chatting on a given topic is mainly used by the students in the evaluation study, and is also the main function of the whole system the students have used. The formal evaluation results indicate the application of CSIEC system in English class can better assist their language learning, e.g., enhance the fluency of English, the confidence on English communications, the interest in English, the mastery of practical expressions, and listening skills. The planned system functions including free chatting and chatting on a given topic, and listening training have been brought into pedagogical play.

The CSIEC has been practically applied since its birth. We continue to improve its interface and content according to the user feedback, either from free Internet users or from the English class students. Such new functions as talk show and adjustable speaking speed of the agent characters are originally suggested by the users. The design, implementation, application and evaluation are not separated, but integrated together. This kind of application-driven research can quickly transfer the user's demands into technical implementation, and new emerging technologies into pedagogical application. It is also consistent with the design-based research theory, which came into birth in 1990s with the goal to fill the gap between the practice and the traditional evaluation research about the integration of technology and education, and to enhance the integration of technology into curriculum and learning efficiency through practice oriented research. It combines the learning environment design and theoretical development, and stresses the research via continuous and upgraded rotation of design, implementation, feedback and analysis [17].

Through the application and evaluation we find currently there are still some user requirements which haven't been fulfilled well, for example, the system's stronger ability of natural language understating and generation, which is the fatal factor influencing the human-computer communication, the lifelike synthesized agent voice and high response speed, which also have been addressed in the users' feedback. Solely in NLP many problems are still hard to be solved, such as the textual ambiguity and entailment [18]. How to overcome these problems with current available technologies is still a great challenge to us.

Acknowledgments

We are grateful for the support to our projects from Konrad-Adenauer-Foundation Germany, Ministry of Education China, Peking University, Education Committee of Capital Beijing, and Korea Foundation for Advanced Studies.

References

1. Ministry of Education China. Annual Educational Statistics 2006. People's Education Press, Beijing (2006)
2. Fryer, L., Carpenter, R.: Emerging Technologies Bots as Language Learning Tools. Language Learning &Technology 10(3), 8–14 (2006)
3. Liang, X.: Trial of Multimedia English Instruction Network. Journal of Hebei Medical University 26(6), 573–574 (2005)
4. Brennan, K.: The Managed Teacher: Emotional Labour, Education, and Technology. Educational Insights 10(2), 55–65 (2006)
5. Weizenbaum, J.: ELIZA–a Computer Program for the Study of Natural Language Communications between Man & Machine. Communications of ACM 9(1), 36–45 (1966)
6. Graesser, A.C., Chipman, P., Haynes, B.C., Olney, A.: AutoTutor: An Intelligent Tutoring System with Mixed-initiative Dialogue. IEEE Trans. on Education 48(4), 612–618 (2005)
7. Seneff, S.: Interactive Computer Aids for Acquiring Proficiency in Mandarin. In: Huo, Q., Ma, B., Chng, E.-S., Li, H. (eds.) ISCSLP 2006. LNCS (LNAI), vol. 4274, pp. 1–12. Springer, Heidelberg (2006)
8. Kerfoot, P., et al.: A Multi-institutional Randomized Controlled Trial of Web-based Teaching to Medical Students. Academic Medicine 81(3), 224–230 (2006)
9. Abu Shawar, B., Atwell, E.: Fostering language learner autonomy via adaptive conversation. In: Proceedings of Corpus Linguistics (2007), `http://corpus.bham.ac.uk/corplingproceedings07/paper/51_Paper.pdf`
10. Kerly, A., Hall, P., Bull, S.: Bringing Chatbots into Education: Towards Natural Language Negotiation of Open Learner Models. Knowledge-Based Systems 20(2), 177–185 (2007)
11. Frege, G.: Begriffsschrift, Eine der Arithmetischen Nachgebildete Formalsprache des Reinen Denkens. Wissenschaftliche Buchgesellschaft, Darmstadt (1879)
12. Wittgenstein, L.: Tractatus logico-philosophicus. Suhrkamp, Frankfurt am Main (1918/21)
13. Jia, J., Ye, Y., Mainzer, K.: NLOMJ–Natural Language Object Model in Java. In: Shi, Z., et al. (eds.) Intelligent Information Processing II, pp. 201–209. Springer, NY (2004)
14. Fellbaum, C.: WordNet: An Electronic Lexical database. MIT Press, Cambridge (1998)
15. Jia, J., Chen, W., Hou, S.: Improving the CSIEC Project through Agent Technology & Adapting It to English Instruction in China. In: Proceedings of ICCE 2006, pp. 47–54 (2006)
16. Jia, J.: The Study of the Application of a Web-Based Chatbot System on the Teaching of Foreign Languages. In: Proceedings of SITE 2004, pp. 1201–1207. AACE, VA (2004)
17. The Design-Based Research Collective. Design-Based Research: An Emerging Paradigm for Educational Inquiry. Educational Researcher, 32(1), 5–8 (2003)
18. Dagan, I., et al. (eds.): MLCW 2005. LNCS (LNAI), vol. 3944. Springer, Heidelberg (2006)

A Strategy for Selecting Super-Peer in P2P and Grid Based Hybrid System

Sheng-Hui Zhao[1,2], Gui-Lin Chen[1], Guo-Xin Wu[2], and Ning Qian[2]

[1] Department of Computer Science & Technology, Chuzhou University, Anhui, China
[2] School of Computer Science & Engineering, Southeast University, Nanjing, China
shzhao@ah.edu.cn

Abstract. With the explosive increase of digital resources, the efficiency of searching resources has become a key problem. Both P2P and Grid are distributed network system, which provide good platforms for storing resources. In the hybrid system merging Grid and P2P, using super-peers can get better effect of searching resources. Combining the characteristics of P2P with Grid, a super-peer selection strategy is proposed. In the experiments, success rate of query resources and query delay are evaluated respectively, as well as the comparison between our selection strategy and random selection method. The results show that super-peer by our selection strategy can improve efficiency of querying resources, decrease query delay and accelerate the querying rate.

Keywords: Sharing, Super-peer, P2P, Grid.

1 Introduction

The rapid development of Internet bring us much convenience, following the explosive datum. Resolving data storage and access is significant.Many techniques try to solve the problem, like as data warehouse, SAN, P2P, Grid, etc. Both P2P and Grid can realize resources sharing and have different advantage and shortage.Therefore, we may combine their merits and establish a convergence system of P2P and Grid.

Napster,Gnutella and KaZaA are all well known P2P system, and Globus Toolkit (GT) [16] has been accepted as the mature software toolkit on deploying Grid, which make it probability combining P2P and Grid. One of GT 4.0 component is Monitoring and Discovery System (MDS). MDS is a suite of web services to monitor and discover resources and services on Grids. A service for collecting above information is Index Service included in MDS. Indexes collect information and publish those information as resource properties. It provides information to clients by a Web Service interface. Clients may query and subscribe resource properties from an Index. Index Service not only saves the local useful datum, but also caches the remote datum, and maintains data updating by lifetime management mechanism. In the large scale Grid, indexes can register to each other in a hierarchical fashion in order to aggregate data at several levels.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 192–199, 2008.

Utilizing MDS of GT4.0, we can monitor and attain nodes'resources information which are divided into dynamic, for example, available disk space, network bandwidth, delay, idle physical memory, free CPU and so on, as well as static information like as operating system name and version, processor type and physical memory size ,etc. Using these properties information, we can compute nodes' capacity.

2 Related Work

Making use of P2P scalability and dynamic, designing resources location and discovery protocol for querying Grid resources to improve query success rate and fault tolerance of system have been already involved into network systems merging P2P and Grid[1] [2] [3] [4] [5] [6]. These hybrid systems can efficiently and reasonably implement resources utilization, which provide a new platform for sharing digital resources. In these hybrid systems, super-peers may manage general nodes concentratedly within certain range and integrate limit resources, in which super-peers collaborate each other to form a decentralized P2P network at a higher layer. B.Yang et al. studied pure P2P network based on super-peer architectures[7]. KaZaA and Gnutella are two typical existed super-peer based P2P system. Furthermore, Montresor et al. proposed opinions on super-peer selection [8] [9] [10] [11] in their systems, but did not do further research on how to select. C.Mastroianni et al. devised Grid information service on the basis of super-peer model[12] [13]. C.Pasquale et al. presented a model for job assignment across the Grid exploiting an underlying super-peer topology[14]. A hybrid and unstructured network model based on P2P and Grid was introduced in [15]. But these models only gave or mentioned the concept of super-peer. Although they have agreed on the opinions of how constructed super-peer architecture according to P2P and Grid, they had not discussed super-peer selection and its form in detail. Apparently, super-peer plays an important role in the network, so its performance will influence the whole network performance directly. How to reasonably select super-peer has been recognized as a new research hot.

This paper designs a super-peer selection algorithm based on capacity. The rest of paper is organized as follows. Our super-peer selection algorithm based on capacity is described in detail in section 3. Section 4 discusses the experimental environments and evaluates the results. Section 5 concludes the paper.

3 A Strategy for Super-Peer Selection

In the network system of this paper, Grid is deployed with GT4.0, and P2P is unstructured similar to Guntella. Grid is constructed with Virtual Organization (VO) for specified application target. Each VO sets an aggregated node(AN), and every node registers to AN as a service when joining the system. All these nodes are Grid nodes. Index Service in AN collects the resources state information of Grid node. All registered Grid nodes probable become super-peer, and each registered service updates its information periodically.

Super-peers are selected from Grid nodes registered to AN, and they connect to each other to form an overlay network at a higher level using P2P mechanism. In a VO, the number of super-peer is limited. A super-peer and its managing subnodes are called as a cluster. Through limiting the size of cluster, it can guarantee that the amount of super-peers is not far too much.

3.1 Normalization of Resource Properties

Supposing there is n nodes in a VO, each node has t properties. We can obtain Matrix E. Each row in E represents a node, while each column represents one of the node properties. $q_{i,j}$ is the Matrix element. Each property'value has different type and range. In order to allow for an uniform measurement of properties, it should be normalized. Then the range of values can be adjusted to a uniform scope $[0, 1]$.

The properties normalization computing method is formula (1).

$$norm_{i,j} = \frac{q_{i,j}}{max(\{q_{i,j}|j = 1, 2, \ldots, t\})} \tag{1}$$

Each property $q_{i,j}$ has a weight w_j in order of precedence.The sum of w_j is 1, that is $\sum_{j=1}^{t} w_j = 1, w_j \in (0, 1)$.

Set, $h_{i,j} = w_j \times norm_{i,j}$;

Then matrix E is converted into N, $h_{i,j}$ is the element of Matrix N.

Set,

$$C_i = \sum_{j=1}^{t} w_j \times h_{i,j}, i = 1, 2, 3, \ldots, n \tag{2}$$

C_i is said as the capacity of node i.

3.2 Description of Super-Peer Selection Strategy

For the sake of describing the algorithm, firstly,we define two sets. The set of super-peers is $SP = \{sp_{i1}, sp_{i2}, \ldots .sp_{ik}\}$. $CA = \{ca_{i1}, ca_{i2}, \ldots .ca_{ik}\}$ represents the set of nodes'capacity registered to AN. After Grid node registering into AN, AN gets its information, such as free CPU, idle memory, available bandwidth, number of current connection and online time 5 properties, and it stores them in the index table. Capacity is computed based on weight using formula (2) after normalizing the value of properties and be put in CA.

Assuming M is the maximal value of cluster size. When the number of nodes connecting to a super-peer, i.e, the number of nodes in a cluster less than M, this super-peer is named as non-saturated. ca_i is the capacity of node i, and NP_{j_conn} represents the number of connecting to a node j, while $|SA|$ represents the number of saturated super-peers. R is the number of subnodes which would be released by super-peer when a new super-peer joins the system.

In order to balance the load, when adding a new super-peer, AN will announce the existed super-peers to release some subnodes and arrange the released subnodes to connect to the new super-peer.

Algorithm SSABC:

Step 1 Initialization: $SP \leftarrow \phi, CA \leftarrow \phi$
Step 2 if first node S_1 joins the network then
  $SP \leftarrow SP \cup \{S_1\}$;
  $CA \leftarrow CA \cup \{ca_{s_1}\}$;
Step 3 if the ith node S_i joins the network then
  $CA \leftarrow CA \cup \{ca_{s_i}\}$;
Step 4 /* if exist super-peer is non-saturated in SP , then S_i connects to the non-saturated super- peer which is the nearest to S_i. Using hop counts computes the distance .*/
  if $\exists NP_{j\_conn} < M \wedge S_j \in SP$ then
    S_i connects to S_j;
    go to Step 3;
  endif
Step 5 /*if super- peers in SP are all saturated, then select a new super-peer.*/
    /* if S_i's capacity is the maximal then select it as super-peer.*/
  for each ca_{s_j} in CA
    if $ca_{s_j} < ca_{s_i}$ then
      $SP \leftarrow SP \cup \{S_i\}$;
      $CA \leftarrow CA - \{ca_{s_i}\}$;
    else
    /* S_i 's capacity is not the maximal, then select a node from CA whose capacity is the maximal, assume it as S_k .*/
      Disconnect S_k from its super-peer by AN announcement;
      $SP \leftarrow SP \cup \{S_k\}$;
      $CA \leftarrow CA - \{ca_{s_k}\}$;
    endif
  endfor
Step 6 /* let every existed super-peers released the number of R nodes ,balancing the load*/
  if S_i or S_k as super peer then
    $R = M - M \times |SA|/(|SA| + 1)$;
    $R \times |SA|$ subnodes connect to S_i or S_k ;
  endif

SSABC algorithm is executed in one VO.

4 Experiments and Evaluation

When a node comes into system and register to AN, it can be as a Grid node, and AN's MDS can find its current resources state. If it is selected as a super-peer, it can deal with user resource application and provide resources locating in itself or cluster to clients . Each super-peer has an index table, which saves its neighbors super-peers' state information of available resources. Neighbors can gain update-registering information each other by periodically announcement.

While a client applies for resources, its node will submit application to its super-peer. If super-peer couldn't satisfy client's resources application, it would search its index table for finding a neighbor super-peer to satisfy client requirement. If it still can't find the target super-peer, the origin super-peer will randomly select a neighbor super-peer on behalf of its client and continue the same query. In the process of query, we set TTL value to limit the query be carried out in VO. When TTL decrease to zero, query would fail.Our experiments simulate the behavior of client querying file resources and other resources such as free CPU, available bandwidth, memory, etc.

4.1 Experimental Environments

We make the simulation program utilizing VC++6.0. The network topology in our experiments adopts the random graph model based on WAXMAN. Making use of BRITE [17], an Internet topology generator, we generate multi-different network topology files aiming at various number of nodes. Simulation parameters and corresponding values used in our analysis are listed in Table 1. Each node

Table 1. Simulation Parameters

Parameter	Values	Parameter	Values
Kinds of Files	100	Online Time	360∼2400minutes
Idle CPU	{1,2,3,4}	Total Nodes	6000 ∼10000
Available bandwidth	10 ∼ 1024Mbps	TTL	2 ∼ 7

bandwidth and current connection nodes are generated in topology files. The number of file resources follows geometric distribution. The number of CPU, memory and online time follow uniform distribution. These values are produced by Matlab random functions. The literature [18] pointed that cluster size in KaZaA system was between 60 and 100, while ultrapeers in Guntella system reached 30 ∼ 40. So we set the proportion (P_{sp}) having super-peers as 1%∼ 5% in one VO. The weight for properties of CPU,bandwidth,memory, online time and the number of connection is 0.25,0.25,0.1,0.2 and 0.2,respectively, .

Query success rate is looked as success frequency divided by total number of times of success. Successful query which means finding requesting resources will produce many results, and each result has a latency. Average query delay is total query latency divided by total query number of times. In order to explain the advantage of capacity selection (CS) super-peer, we compare it with random selection (RS) super-peer.

4.2 Experimental Results

The total number of nodes are 8000 in Figure 1-2. The two figures are all capacity selection super-peer itself comparisons,where query resources is file type. P_{sp} in Figures 1-2 is 3%. To express conveniently, P_{sp} in Figure 3-4 is expressed by SP=%, not P_{sp} =%. Total number of times of query is 1000.

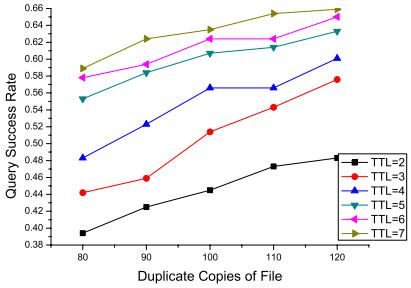


Fig. 1. File Duplicate Copies versus Query Success Rate

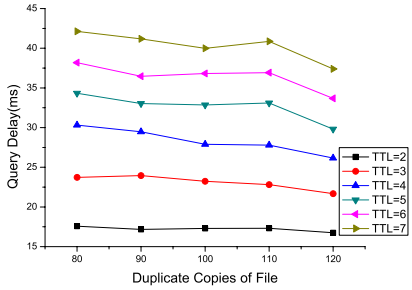


Fig. 2. File Duplicate Copies versus Query Delay

In Figure 1, while TTL is fixed, success rate appears ascending trend as the duplicate copies of file rising from 80 to 120. File duplicate copies versus success rate nearly presents linear relation. Figure 2 demonstrates that average query delay doesn't change obviously for a fixed value of TTL, and the curves are relatively steady. When TTL increases, average query delay rise from 17ms to 45ms. Figures 1-2 show that our super-peer selection strategy is steady and reasonable.

Success rate of between CS and RS with different TTL are shown in Figure 3 when P_{sp} is changed from 1% to 5%, where query resources is expressed with capacity for the purpose of conveniences and simplicity. Capacity values is generated by random function in program. This method doesn't effect performance analysis, because comparison of query success rate is based on same capacity,TTL and P_{sp}. By all appearance, for a fixed number of super-peers, the success rate in CS is higher than RS. However, the number of super-peers hardly makes any difference to success rate under random selection.

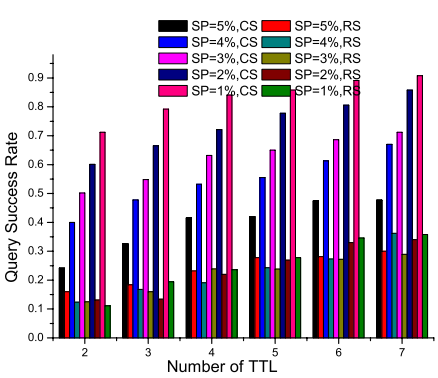


Fig. 3. Query Success Rate between CS and RS with Different TTL

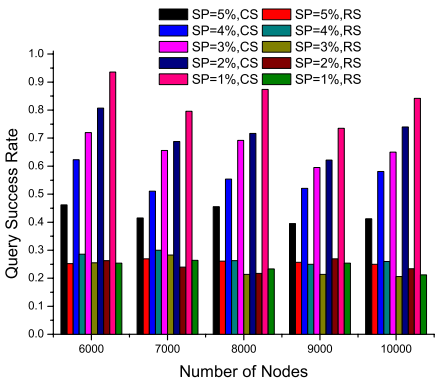


Fig. 4. Query Success Rate between CS and RS with Different Number of Nodes

Figure 4 shows the comparison of success rate between CS and RS with different number of nodes. From figure 4, it is observed that the success rate of both CS and RS are all not clearly change at the same P_{sp}. On the condition of same nodes, success rate of CS is still increasing with cluster size, while success rate of RS doesn't make better. It can be seen from Figure 3 and Figure 4, increasing nodes doesn't influence query success rate, which explains the capacity selection super-peers having better stability,and shows our capacity selection super-peer is priority to random selection super-peer and advances the query efficiency.

5 Conclusions

Super-peer model based network system has widely known, and how to select high performance super-peers is a key problem in these systems. However, research on selecting super-peer mostly has been done in pure P2P systems. Utilizing MDS4.0 index service to discover dynamic useful resources and using resources properties to compute nodes'capacity, this paper presents a strategy for selecting super-peer according to nodes'capacity in P2P and Grid based hybrid system. Experimental results demonstrate that success rate of capacity selection is obviously higher than that of random selection. Therefore, selecting higher capacity nodes as super-peers in the convergence system of P2P and Grid improves query success rate and decreases query time, which provides a better platform for sharing digital resources.

Acknowledgement

The work presented in this paper is supported by the Natural Foundation of Anhui Provincial Education Department (No.2006KJ041B and No.KJ2007B073)

References

1. Puppin, D., Moncelli, S., Baraglia, R., Tonellotto, N., Silvestri, F.: A Peer-to-peer Information Service for the Grid. In: Proc. of International Conference on Broadband Networks, San Jose, CA, USA (2004)
2. Andrade, N., Costa, L., Germ'oglio, G., Cirne, W.: Peer-to-peer grid computing with the OurGrid Community. In: 23rd Brazilian Symposium on Computer Networks - IV Special Tools Session (May 2005)
3. Amoretti, M., Reggiani, M., Zanichelli, F., Conte, G.: SP2A: Enabling Service-Oriented Grids using a Peer-to-Peer Approach. In: 14th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprise (WETICE 2005), pp. 301–304 (2005)
4. Uppuluri, P., Jabisetti, N., Joshi, U., Lee, Y.: P2P Grid: Service Oriented Framework for Distributed Resource Management. In: IEEE International Conference on Services Computing (SCC 2005), pp. 347–350 (2005)
5. Iamnitchi, A., Foster, I.: A peer-to-peer Approach To Resource Location In Grid Environments. In: Grid Resource Management. Kluwer Publishing, Dordrecht (2003)

6. Talia, D., Trunfio, P.: A P2P Grid Services-Based Protocol: Design and Evaluation. In: Danelutto, M., Vanneschi, M., Laforenza, D. (eds.) Euro-Par 2004. LNCS, vol. 3149, pp. 1022–1031. Springer, Heidelberg (2004)
7. Beverly Yang, B., Garcia-Molina, H.: Designing a Super-Peer Network. In: 19th Int'l Conf. on Data Engineering. IEEE Computer Society Press, Los Alamitos (2003)
8. Montresor, A.: A Robust Protocol for Building Superpeer Overlay Topologies. In: Proc. Of the International Conference on Peer-to-Peer Computing, Zurich, Switzerland (2004)
9. Li, J., Vuong, S.: An Efficient Clustered Architecture for P2P Networks, In: 18th International Conference on Advanced Information Networking and Applications (AINA 2004), vol. 1, p. 278 (2004)
10. Lo, V., Zhou, D., Liu, Y., GauthierDickey, C., Li, J.: Scalable Supernode Selection in Peer-to-Peer Overlay Networks. In: Second International Workshop on Hot Topics in Peer-to-Peer Systems, pp. 18–27 (2005)
11. Min, S.-H., Holliday, J., Cho, D.-S.: Optimal Super-peer Selection for Large-scale P2P System. In: 2006 International Conference on Hybrid Information Technology - Vol2 (ICHIT 2006), pp. 588–593 (2006)
12. Mastroianni, C., Talia, D., Verta, O.: A Super-Peer Model for Building Resource Discovery Services in Grids: Design and Simulation Analysis. In: Sloot, P.M.A., Hoekstra, A.G., Priol, T., Reinefeld, A., Bubak, M. (eds.) EGC 2005. LNCS, vol. 3470, pp. 132–143. Springer, Heidelberg (2005)
13. Puppin, D., Moncelli, S., Baraglia, R., Tonelotto, N., Silvestri, F.: A Grid Information Service Based on Peer-to-Peer. In: Cunha, J.C., Medeiros, P.D. (eds.) Euro-Par 2005. LNCS, vol. 3648, pp. 454–464. Springer, Heidelberg (2005)
14. Cozza, P., Talia, D., Mastroianni, C., Taylor, I.: A Super-Peer Model for Multiple Job Submission on a Grid. Technical report, TR-0067, Institute on Knowledge and Data Management, Institute on Grid Systems, Tools and Environments, CoreGRID - Network of Excellence (January 2007)
15. Guilin, C., Shenghui, Z., Zhengfeng, H.: A model of integrating P2P technology and grid technology. Journal of Hefei University of Technology(Natural Science) 30(6) (2007)
16. GT4.0, http://www.globus.org
17. BRITE, http://www.cs.bu.edu/brite/
18. Qiong, L., Peng, X., Hai-Tao, Y., Yun, P.: Research on Measurement of Peer-to-Peer File Sharing System. Journal of Software 17(10) (October 2006)

Personal Knowledge Management in E-Learning Era

Weichao Li[1,2] and Yong Liu[1]

[1] Information Science Department, Zhengzhou Institute of Aeronautical Industry Management, Zhengzhou 450015, China
[2] Information Management Department of Nanjing University, Nanjing 210093, China

Abstract. In this paper, the author first introduces some relative theories of E-Learning and personal knowledge management(PKM), including the concept, content, objectives and their common ground, then puts forwards the requirements to personal knowledge management and how to use personal knowledge management tools such as iSpace Desktop, iNota, Mybase in E-Learning environment to improve personal knowledge literacy.

Keywords: E-Learning, Personal knowledge management, Personal knowledge management tool, Knowledge literacy.

1 E-Learning and Personal Knowledge Management

1.1 Introduction to E-Learning

The new management master of the United States, Peter M. Senge has said "If a person can't update what he has learned at the race of 7 percent annually, then he will not be able to adapt to social change." [1]. The knowledge economic era, which is characterized by faster and faster speed of knowledge conversion and technical updates, is a new era of learning which makes learning social and lifelong. E-Learning is derived from E-commence, and different organizations have different definitions, among which the definition made by the U.S. Department of Education is: E-Learning is the learning and teaching activities mainly conducted through the Internet, which make full use of the learning environment with new communication mechanism and rich resources provided by modern information technology, so as to achieve a brand-new mode of learning.

The report Knowledge-oriented society Based on E-Learning, published by Advanced Learning Infrastructure Consortium of Japan in March 2003, pointed out that E-Learning is a new-type learning environment, an active learning activity conducted by learner through applying information technology and communication networks. E-Learning has many features, such as customized courses, active and interactive learning, learning outcomes and learning process easy-controlled, learning anytime & anywhere and for anyone, transmission to the scattered learners, transmission quickly and timely, learning content easy-archived and easy-reuse, etc. [2]. The term E-Learning refers to the employment of new technology for learning purposes, and two key factors in any E-Learning system are learning content creation and delivery [3]. The term E-Learning refers to the digital learning process in digital learning environment, under

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 200–205, 2008.

which the learner makes use of digital learning resources, digital learning environment, digital learning resources and digital learning method are the three basic elements of E-Learning [4].

1.2 E-Learning Content and Objectives

The term E-Learning which stresses on the philosophy of "interactive learning", "life-long learning" and "learner-centered", is typically Internet interactive learning environment. Learning in the E-Learning environment, every learner can freely select appropriate resources from Internet according to his/her own characteristics and can study in terms of his/her manner and speed.

The currently E-Learning model - "Five-A learning model", is that learner can freely select appropriate resources from Internet according to his/her own characteristics to choose Anywhere, Anytime, Anything, Anyone, Ability mode of learning [5]. The characteristics of E-Learning can be roughly summed up as flexible, easy-to-get and convenient.

E-Learning is aimed at generally improving the younger generation in basic learning skills, information quality, innovative ability, interpersonal communication & cooperation and practical ability so as to cultivate a large number of creative talents in 21st century.

1.3 What Is PKM

For the concept of personal knowledge management, there is still no uniform definition. Viewing from abroad, the broad definition by Paul, the United States is: PKM should be seen as a set of problem-solving skills and methods in level of both logical concept and practical operation [6]. Frand and Hixon believe that PKM refers to a strategy and process to expand personal knowledge, during which individuals organize and concentrate their important information as a part of their own knowledge, and transfer scattered fragments of information into systemic application information. In addition, they believe that PKM also includes the expansion of personal knowledge and the conversion from personal tacit knowledge to explicit knowledge [7]. In domestic research, Dechao KONG believes that the PKM includes three meanings: First, managing personal knowledge gained; Second, acquiring new knowledge through various channels, learning from and drawing on the experience and strong points of others to make up for their own deficiencies in thinking and knowledge, so as to constantly construct their own knowledge characteristics; Third, achieving dominant change of tacit knowledge and stimulating innovation of new knowledge by applying their mastered knowledge and the long-standing views and ideas, and combining with other people's ideological essence and disposing draft [8].

PKM is the management of knowledge resources to achieve personal goals. Based on PKM, it emphasizes on recording and digging personal tacit knowledge, accessing knowledge resources necessary for work and learning, promoting knowledge resources orderly and in self-organization, and promoting reproduction and reuse of personal knowledge. PKM mainly includes the management of basic personal knowledge database and the management of personal thinking database. The former includes personal communications management, personal time management, personal work

management, personal learning management, personal network resource management and personal files management, etc. The latter includes personal knowledge digging, personal intellectual property management, personal knowledge marketing, etc. [9].

1.4 PKM Skills

Professor Dorsey regards that PKM can include and define the following seven core skills: retrieving information skills, evaluating information skills, organizing information skills, analyzing information skills, presenting information skills, securing information skills and collaborating around information skills.

1.5 Common Ground of KM and E-Learning

The common ground is reflected in the following aspects:

(1) Collaborating / cooperating; The collaboration is a key factor in the E-Learning environment under which both the distribution of learners and transmission of knowledge have a characteristic of separation of time and space. If the learner wants to solve problems encountered during learning, it is very necessary for he/she to get collaboration from partner and a guide from companions Student agent. According to Chris Christiansen and other people's viewpoints, in the E-Learning environment, collaboration can overcome two main issues of distance learning: adapting to distance learning and establishing distance learning community [10].

(2) Trusting and knowledge sharing; Technology can only make knowledge sharing easier, but mutual trust among learners can make knowledge sharing possible. If the individual does not trust other people’s knowledge, or does not believe that other will contribute their knowledge, the team is not very effective.

(3) Shared understanding; As for effective knowledge sharing, learners must have the same understanding in the process of communication. In the E-Learning environment, shared understanding is crucial to learners during the learning process, which can deepen individual learning process into organizing behavior.

(4) Information technology; Both fully rely on information technology, such as computer technology, network technology, communications technology, multimedia technology, etc..

(5) Virtual community; In the view of knowledge management, community is a very important place for knowledge collaborating and sharing. The community is dynamic and rich learning model in knowledge creation and sharing.

2 Requirements of E-Learning to PKM

In knowledge economic era, the output of information and knowledge are growing exponentially, and anyone can gain massive information from Internet, but he/she does not know how to extract useful information from the ocean of information and then effectively transform them into their own knowledge. That "We live in the ocean of information, but put up with the thirst of knowledge" is a vivid portrayal of the situation [11]. Faster and faster the volume of information increases, the more and more complex forms of knowledge and the faster speed of knowledge updating all make PKM a hot topic in today's society.

If a person wants to obtain the survival and development capacity and to keep pace with the rhythm, he must become a personal knowledge manager to enhance the capacity of creating and applying knowledge. In addition, during the process of PKM, such as the discovery of knowledge, access of knowledge, storage of knowledge, share of knowledge, pervasion of knowledge, application of knowledge and creation of knowledge, and so on, involving methods, the exertion of strategy and means are involved. So how to manage knowledge is an important lifelong learning content. Therefore, it is very important for E-Learning to provide learners with corresponding tools to help study. In many helpful tools of learning [12], during the learning process if we can apply knowledge management technology to promote the conversion between tacit and explicit knowledge, to promote the application of knowledge management technology and tools and to promote knowledge innovation and sharing, it will be more helpful to enhance the effectiveness of lifelong learning and to promote individual lifelong learning more effective.

3 Applying PKM Tool in E-Learning

3.1 PKM tools' Category

Knowledge has become the most important personal core resources in this era of lifelong learning. PKM is bound to produce an extensive and far-reaching impact on education, and corresponding PKM tools will be fully applied in E-Learning. PKM tools generally support several knowledge management processes, i.e. coding / describing, classifying / indexing, searching / filtering, sharing / disseminating and knowledge innovation [13]. PKM tools are different according to their core functions and can be classified as follows:

(1) Information capture / sharing tool

Such tools drag information (text, graphics, charts, links, etc.) from the Web page or document to form new different type document. When information sources add new content or information updates, the tool can inform users, and share the information with others. The tool has a series of products of Entopia: Quantum Collect, Quantum Collaborate and Quantum Capitalize, Web2one, Organizer, and so on.

(2) Encoding / describing tool

Such visualization tools can capture, organize and display (in the form of link and concept) new concepts, and provide with functions of searching, enlarging, expanding and navigating, etc.. Such tools are Mind map, Brain, etc..

(3) Search / indexing tool

Such tools index local and network drives, support keyword, full-text and natural language, Boolean expressions, etc.. For example EnFish Find can look for and classify information according to correlativity, names, documents, E-mail and URL, and then provide online communities to support collaboration among individuals.

(4) Meta Search tool

Meta search engine will send questions to different database of several search engines to search, and provide a single integrated and graded list. Such tools are Scarch.com, Dogpile.com, Mamma.com, etc..

(5) Reasoning links Wizard

Such tools will pop up with links related with the contents of documents, provide online dictionary and lexicon, and support information access and sharing. Such tools are Atomica Personal, DB / TextWorks, etc..

(6) Collaboration / Synchronization tool

Through the form of questioning / answering, discussing and ideas sharing, the knowledge is shared among the group of mutual interest in a theme. Individuals can confirm theme and sub-theme, subscribe and unsubscribe at any time, accept the new post with a choice, regularly produce summary report of the theme. More famous tool of this kind is Yahoogroup.com.

(7) Learning tools

Such tools are designed to help individuals control the learning process, collect and prepare courseware, train & guide, track personal capacity changes. Such tool is Digital Learning System (DLS) of BrainX.

3.2 Examples of PKM Tools

(1)PKM System (iSpace Desktop V1.1.2)

ISpace Desktop is an integrated PKM system with personal information management, knowledge management and communication management as its basic task helping individual effectively manage personal information related with work, study, daily life and social relations, including: address book management, document management, schedule management, blog readers, pages Internet browser, etc. [14].

(2) PKM Tools (iNota)

This is a PKM tool for editing, which can gain text or graphics in form of drag or clip, and classify and manage information in structure of tree. It notes the information in detail, automatically transfers it into XML document used as network resource, and then makes up and classifies, establishes personal directory and PKM system by method of noting by emphasis and adding contents, and improves efficiency of information management and knowledge absorption. Its main features are as follows: simple interface, clear information classification, detailed information endorsement, concise search, easy data storage, automatic files transfer, and so on [15].

(3) Documentation Resource Management Software (Mybase V5.2)

Mybase is a universal data compression manager which has powerful functions and can freely customize the format and level relation. It can be used to manage a variety of information, such as: various types of documents, disk documents, data, business cards, events, essence downloaded and information collected. Even documents without any rules can be managed methodically. If you are good at managing information, Mybase will become a handy tool for you; if you are not good at managing information, Mybase will help you improve information management capabilities [16].

4 Conclusion

In the current knowledge economic era, knowledge has become an extremely important factor for personal survival and development. The individual must skillfully master PKM skills so as to gain the survival and development ability and become a constantly

growing individual to realize his-self value. Individual learning mode is impacted by information technology revolution and has developed to E-Learning phase, and at the same time individual also changes the E-Learning through mastering PKM tools to enhance personal knowledge literacy.

References

1. Lai, C.S., Lu, Z.H.: Initial Discussion on Personal Knowledge Management Method. Jinzhou Medical College Journal 21, 76–78 (2000)
2. Advanced Learning Infrastructure Consortium: Knowledge-oriented Society Based on E-Learning (2003)
3. Henry, P.: E-learning Technology, Content and Services. Education Training 43, 249–255 (2001)
4. Li, K.D.: Digital Learning (1)-Core of Information Technology and Curriculum Integration. Education Research 100, 46–49 (2001)
5. Luo, Z.M., Chen, C.J.: Research on Definition and Characteristics of E-learning. Foreign Language Teaching 26, 60–64 (2005)
6. Dorsey, P.A.: What is PKM? http://www.millikin.edu/webmaster/seminar/PKMnet/whatispkm.htm
7. Personal Knowledge Management: Who, What, Why, When, Where, How?, http://www.anderson.ucla.edu/faculty/jason.frand/researcher/speeches/educom98pkm/sld001.htm
8. Kong, D.C.: Personal Knowledge Management. Library Development 3, 17–18 (2003)
9. Ke, P.: Introduction to Information Management, 2nd edn. Science Press, Beijing (2007)
10. Barth, S.: Personal Toolkit: A Framework for Personal Knowledge Management Tools, http://www.kmworld.com/publications/magazine/index.cfm?action=readarticle&Article_ID=1406&Publication_ID=83
11. Chen, L.X.: Personal Knowledge Management. Information Science 23, 1072–1075 (2005)
12. Zhong, Z.M.: The Research of Supporting Platform of Asynchronous Q & A for E-learning. J. of Zhengzhou Univ (Nat. Sci. Ed.) 39, 55–59 (2007)
13. Gan, Y.C.: Personal Knowledge Management in E-Learning Environment. China Educational Technology 197, 20–24 (2003)
14. http://soft.ccw.com.cn/download (2007.12.06)
15. http://www.360doc.com/showweb/0/0/204026.aspx (2007.12.06)
16. http://www.skycn.com/soft/4039.html (2007.12.06)

Teaching Machine Learning to Design Students

Bram van der Vlist, Rick van de Westelaken, Christoph Bartneck, Jun Hu, Rene Ahn, Emilia Barakova, Frank Delbressine, and Loe Feijs

Department of Industrial Design
Eindhoven University of Technology
Den Dolech 2, 5600MB Eindhoven, The Netherlands
{b.j.j.v.d.vlist, h.f.m.v.d.westelaken}@student.tue.nl,
{c.bartneck, j.hu, r.m.c.ahn, e.i.barakova, f.l.m.delbressine, l.m.g.feijs}@tue.nl

Abstract. Machine learning is a key technology to design and create intelligent systems, products, and related services. Like many other design departments, we are faced with the challenge to teach machine learning to design students, who often do not have an inherent affinity towards technology. We successfully used the Embodied Intelligence method to teach machine learning to our students. By embodying the learning system into the Lego Mindstorm NXT platform we provide the student with a tangible tool to understand and interact with a learning system. The resulting behavior of the tangible machines in combination with the positive associations with the Lego system motivated all the students. The students with less technology affinity successfully completed the course, while the students with more technology affinity excelled towards solving advanced problems. We believe that our experiences may inform and guide other teachers that intend to teach machine learning, or other computer science related topics, to design students.

Keywords: teaching, machine learning, design, lego.

1 Introduction

The Department of Industrial Design at the Eindhoven University of Technology prepares students for a new type of engineering discipline: design and creation of intelligent systems, products, and related services. These systems, products and services require to be adapted to the user and thereby provide a new experience. In the framework of our Masters program, we offer a course that familiarizes students with a number of powerful conceptual and intellectual tools to understand and create adaptive behavior at a system level.

System level thinking has had and still has an enormous impact upon the development of technology. When working at a system level one does not study individual component behavior, such as Ohm's law for an electrical component; instead one addresses bigger questions such as the stability of the feedback loops, information throughput, or learning capacity. The learning objective includes classical control, reinforcement learning and adaptive control and pattern recognition. The context of

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 206–217, 2008.

Lego is chosen because it is already an example of a system. The project's creative goal is to make a leap forward, extending the scope of the existing system such that adaptive behavior becomes the central theme.

Like many other design departments, we are facing the challenge of teaching the mathematical foundation of machine learning to students that are neither mathematicians nor computer scientists. As a general framework we use a competency based learning model [1-3] that focuses on complex behavior and gives equal weight to knowledge, skills and attitudes. The knowledge, skills and attitudes are integrated already during learning (not afterwards, when the student has become active as a professional). The competencies that students acquire during the learning process are made visible in an individual portfolio. Competency based learning requires a powerful and rich learning environment. This learning model applies particularly well for the profession of industrial designer, where pure knowledge is not enough. The student has to learn how to develop contexts of use, how to actively explore concepts, how to evaluate alternative solutions, how to bring new artifacts into the world, in other words, how to design. Although this appears to be well-accepted for traditional industrial design, where the material form of things is the central theme, it was not a priori obvious whether this learning model could be used for those aspects of industrial design that overlap with computer science. Note that in the near future even the most mundane everyday objects will have embedded electronics or computers and hence the design profession is changing accordingly.

Most of the students in our department do not have an inherent affinity towards technology. They do not build up in-depth knowledge of programming or math.

One of the difficulties in teaching machine-learning is that its theory is abstract. The process and the results of the machine learning are only available inside a computer program. Design students are used to create and work with artifacts in the real world, not with mathematical formulas. This abstraction level inhibits their understanding and makes it difficult for them not only to reproduce relevant knowledge, but also to apply and extend it.

We therefore created a new teaching method to better support students in their learning of machine learning. Our new method involves the usage of embodied intelligence; transferring the abstract theory into a more hands-on experience. We will elaborate on the structure of the course, the materials used, and two concrete case studies. Our method is not limited to machine learning, but can be used to teach many other aspect of computer science to design students. We believe that our insights may inspire and guide other teachers to create better courses for their design students.

2 Structure of the Course

The course's first two weeks are theory oriented. A week during this phase typically consists of two days of theory at the start, followed by three days of practice with an intermediate moment of contact between students and teachers to discuss their progress and to answer specific questions. In these two weeks the students work on very specific methods and principles. During the third and forth week the students are invited to demonstrate their understanding of the theory through something that they create. The teachers encourage depth, through additional theory, tools and methods.

We will now provide a more in depth view on the content of the course, but we would like to emphasize that the method may also be applied to teach different aspects of computer science. In our specific course, the goal is to teach the principles of reinforcement learning and supervised learning to design students.

2.1 Embodied Intelligence

We selected Q-learning and Neural Networks as basic examples of reinforcement learning and supervised learning. We embedded this form of intelligence into a real body: the Lego Mindstorms NXT. Lego Mindstorms is an excellent prototyping platform [4] for creating embodied intelligence. The platform features an NXT brick that includes a microprocessor capable of running a Java virtual machine. It comes packaged with several plug-and-play sensors and actuators and is, by definition, compatible with the Lego brick system. Prototypes can be built with click-and-connect ease, which allows students to focus on the implementation of the software.

Traditionally, machine learning [5] is demonstrated through a computer program that does not only have to perform the learning, but which also has to simulate the environment on which the input for the learning model is based. By using an embodiment, such as the NXT, the sensory input does not longer need to be simulated. The learning program receives its input directly through the attached sensors that react to the stimuli that are already available in the real world [6]. The learning system could, for example, try to learn from the light sensor that is mounted on the bottom of a robotic car. The goal of such a learning program would be to learn how to follow a black line on the ground. The real world can offer a richness that would be difficult to simulate. In addition, the embodiment allows the students to easily explore the influence of the various variables. This simplifies and enriches the process of understanding the meaning of variables in an algorithm, as one can observe the effects of changing these variables in terms of behavioral changes of the embodiment.

2.2 Participants

The participants of our course are all industrial design master students, who can be classified into two types. The first group consists of students who have a certain affinity with technology. These students like to explore technological principles that are new to them. They have a good understanding about a wide range of technologies and their applications. They also have considerable programming skills, with JAVA as solid basis. This group of students is usually the smaller of the two groups and teaching them machine learning is easier. They might even be satisfied with the traditional non-embodied method, but using the Lego NXT platform considerably increases their motivation.

The second group of students can be described as students who do not have an inherent affinity with technology. They have a limited understanding of technological principles and master programming only up to a basic level. Teaching these students machine learning is the true challenge. It still needs to be acknowledged that students of either type are not mathematicians or computer scientists. These students are used to the creative creation of artifacts and not to formulas and algorithms. The teaching method needs to adapt to these characteristics.

3 Material

For an embodied intelligence course, software and equipment is necessary. While the software is available for free, the hardware does require a certain budget. The basic Lego Mindstorms Education NXT set is currently available for 285 Euro. Our practical experience shows that one set can be shared by a maximum of two students. We will now discuss the required hardware and software in more detail.

3.1 Hardware

The NXT brick is part of the Lego Mindstorms set. The NXT is an embedded system with a plastic casing compatible with the Lego brick system. This way it can easily be integrated into a Lego construction that may also contain the sensors and actuators [7]. Using Lego saves a lot of time in constructing mechanical components compared to other methods. An educational version is available that includes the useful rechargeable battery, a power supply and a storage box. The NXT specifications are:

- Atmel 32-bit ARM main processor (256 Kb flash, 64 Kb RAM, 48 MHz)
- Atmel 8-bit AVR Co-processor (4 Kb flash, 512 Byte RAM, 8 MHz)
- Bluetooth wireless communication (CSR BlueCoreTM 4 v2.0 +EDR System)
- USB 2.0 communication (Full speed port 12 Mbit/s)
- 4 input ports: 6-wire interface supporting both digital and analog interface
- 1 high speed port, IEC 61158 Type 4/EN 50170 compliant
- 3 output ports: 6-wire interface supporting input from encoders
- Display: 100 x 64 pixel LCD black & white graphical display
- Loudspeaker: Sound output channel with 8-bit resolution (Supporting a sample rate of 2-16 KHz)
- 4 button user-interface
- Power source: 6 AA batteries or rechargeable Lithium-Ion battery.

Lego has developed a number of sensors and actuators as part of the Lego Mindstorms set. All these sensors are compatible with the Lego brick system. The basic Lego NXT Education set contains the following sensors and actuators:

- Touch sensor – detects when it is being pressed by something and when it is released again.
- Sound sensor – detects both decibels [dB] and adjusted decibel [dBA].
- Light sensor – reads the light intensity in a room and measure the light intensity of colored surfaces.
- Ultrasonic sensor – measure distances from 0 to 255 centimeters with a precision of +/- 3 cm.
- Servo motor with build-in rotation sensor.

As result of the success of the Lego Mindstorms, other companies developed additional sensors and actuators. Some of these companies, such as HiTechnic Products and Mindsensors.com, provide sensors for the NXT platform such as IR Link Sensor, Gyro Sensor, IR Seeker Sensor, Compass Sensor, Color Sensor, Acceleration / Tilt Sensor, Magnetic Compass, and Pneumatic Pressure Sensor. In addition to the Lego

NXT set, a standard computer is needed to write the programs. The programs are then uploaded to the NXT using either USB or Bluetooth.

3.2 Software

Three software components are necessary for this course. All of them are available for free and they replace the original Lego software. Lego's own software development tool is targeted at children and hence does not offer the flexibility and extendibility required for a university course. An extensive tutorial on how to install the components is available at: http://www.bartneck.de/work/education/masterClassLego/javaInstallNXT/. We will now describe the components in detail.

Java is a platform independent, object-oriented programming language (http://www.sun.com/java/). The language derives much of its syntax from C and C++ but has a simpler object model and fewer low-level facilities. Java applications are typically compiled to bytecode, which can run on any Java virtual machine (JVM) regardless of computer architecture. It is a popular language for embedded systems, such as micro controllers and mobile phones and also the Lego NXT is capable of execute Java programs.

It is advisable to use an integrated development environment (IDE) to write Java programs. Eclipse is the powerful and widely used IDE that offers excellent support for Java and the Lego NXT. Eclipse itself is written in Java and its installation is particularly easy.

To enable the Lego NXT to execute Java programs, its original firmware needs to be replaced with the open source leJOS firmware [8]. The old firmware can be reinstalled at any point. Conveniently, leJOS includes a Java Virtual Machine so that no further software installations on the NXT are necessary to execute Java programs. The leJOS Java library is an extension to the standard Java and enables Java programs to use the platform specific features of the NXT, such as sensors and actuators.

The Java Object Oriented Neural Engine (Joone) is an application that allows users to build, test and train neural networks (http://www.joone.org). It features a convenient graphical user interface. Neural networks trained with Joone can be exported and called from any external Java program. It can therefore easily be integrated into more general Java programs, such as Java programs for the Lego NXT.

4 Case Study 1: Reinforcement Learning with the Crawler and Johnny Q

During one week, the students mounted the NXT brick on wheels and gave it an arm with two Lego NXT electronic motors, creating the crawler (see Figure 1). This crawler has wheels (not driven) to allow free forward and backward movement. In order to move itself, the crawler can only use its arm, which has two joints under motor control. The Crawler has sensors to measure the angle of the joints of the arm and also one distance sensor that "sees" the distance from a wall or another reference object. The NXT brick was programmed in Java to execute the reinforcement learning algorithm (Q-learning). It is positively rewarded if it moves forward and negatively rewarded if it moves backwards. It explores its possibilities and learns how it should move to accumulate a maximal reward. The Crawler starts with seemingly random

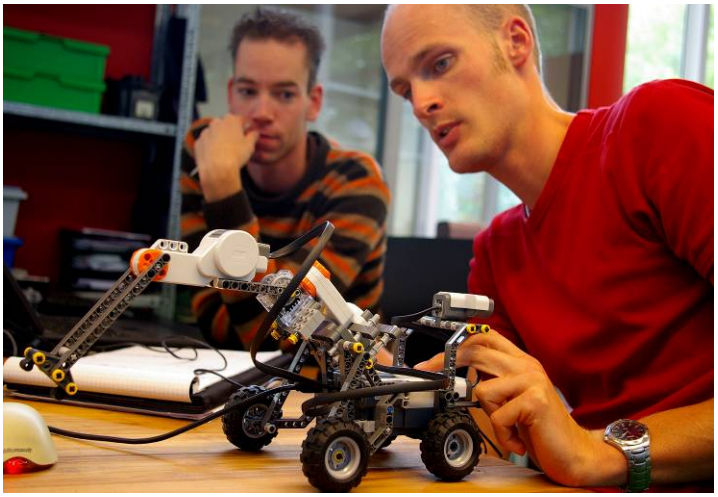


Fig. 1. The Crawler

movements, but after a few minutes it really finds a kind of rhythm allowing it to move the arm and thereby move itself efficiently forward.

A second robot that was built by different students during this week was Johnny Q (see Figure 2). It has wheels and left-right motor drives to move forward, backward, rotate left and right. Johnny Q measures the brightness of the floor and "sees" the distance from a wall or reference object. Inside is an NXT control brick, an embedded processor programmed in Java to execute the reinforcement learning algorithm (Q-learning). The reward is being tapped on the shoulder; a simple button serves to count touches. Johnny Q learns by being trained. Depending on what the human user does or does not reward, Johnny Q learns behaviors, such as turning away from a dark spot, or running backwards near an obstacle. But it can also learn the opposite behavior, bumping against the wall. It explores its possibilities and learns how to accumulate maximal rewards. The observer engages in a training session, teaching tricks and little games, much like training a dog. Usually this algorithm is demonstrated through screen demos but here the potential of embodied learning is visible in a truly embodied model. From a semantic point of view, it is interesting to sculpt the behavior which (of course) requires some patience. Johnny Q will gradually forget although desired behavior can be maintained through continued training.

4.1 Q-Learning Theory

We will now discuss the Q-learning theory in more detail to enable the reader to form a better judgment of the difficulty that the students were able to overcome during one week by using our teaching method. Q-learning is a common and well known reinforcement learning algorithm [9]. Reinforcement learning is a method that allows a machine to learn behavior through receiving rewards and punishments. When a machine performs an action in a certain state it can get a positive reward, negative reward (punishment) or no reward. These rewards reflect the design and goal of the machine. The Q-learning algorithm [10] works by constructing an action-value

Fig. 2. Johnny Q

function that gives an estimate of the expected value *Q(s,a)* (the total award that may eventually be accumulated) when taking a given action *"a"* in a given state *"s"*. Through experience the machine achieves better and better estimates of the action values. The behavior of the machine is given in terms of a policy. The policy determines the probability that the machine will take a certain action in a certain state. An important dilemma in determining the policy is whether the machine should exploit its knowledge and choose the actions that lead to the biggest reward or that it should explore new actions in certain states to discover better ways to retrieve even more rewards later on.

The strength of Q-learning is that it will adapt to its environment without knowing it and without being programmed. Q-learning, as well as other reinforcement learning principles, works because it tries to optimize a given reward. Q-learning requires a finite set of environment states, a fixed set of actions, and a reward function:

$$\pi(s) : A(s) \rightarrow [0,1]$$

$$\forall_s \sum_{a \in A(s)} \pi(s,a) = 1$$

$$Q^{\pi}(s,a) = E_{\pi}\left\{ \sum_{k=0}^{\infty} \gamma^k \tau_{t+k+1} \mid s_t = s, a_t = a \right\}$$

$$Q^{*}(s,a) = \max_{\pi} Q^{\pi}(s,a)$$

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha\left[r_{t+1} + \gamma \max_a Q(s_{t+1},a) - Q(s_t,a_t) \right]$$

$$\pi(s,a) = \begin{cases} \dfrac{\int}{|A(s)|} & \text{if } a \neq \arg\max_{a'} Q(s,a') \\ 1 - \int + \dfrac{\int}{|A(s)|} & \text{if } a = \arg\max_{a'} Q(s,a') \end{cases}$$

We write ε for the exploration factor, γ for discounting factor, α for learning factor, π for policy (ε-greedy), *s* for state, *a* for action, *t* for (discrete) time, $A(s)$ for action set, *Q(s,a)* for expected return in state *s* after action *a*, under current policy, *Q*(s,a)* for expected return in state *s* after action *a*, under optimal policy, and E for expectation.

5 Case Study 2: Voice Command Using Supervised Learning

The students applied their knowledge of neural networks, which is one flavour of supervised learning, to implement a simple speech recognition application. It took them one week to explore the operational principles of a basic neural network and to use this knowledge to design the application. For this application the Lego NXT Sound Sensor and the NXT brick were used to get the desired speech input. The Lego sound sensor is an envelope detector that measures the change in volume (amplitude of the sound signal) over time and not a real microphone. However, the envelope was sufficient input to build a recognition application that could distinguish between the words "Biertje" (beer) and "Champagne" (Champaign) by recognizing the difference in the word's envelopes.

The NXT with help of its microphone recorded the words into sound samples that were then transferred via a Bluetooth connection to the computer. To make sure that the recognition was based on the difference in volume over time and not on the duration of the word, the length of both sound samples was equalized during the preprocessing. The sound samples were fed as input to the neural network that was created using Joone. The resulting output was then communicated back to the NXT that printed the results on its screen.

During the last two weeks of the course, the students were encouraged to create an extension pack for the Lego Mindstorms NXT set. These extension packs should empower other users in the Lego community to easily extend their Lego inventions far beyond what is possible with standard Lego. Two students decided to extend the neural network application that was build in the previous week.

The goal was to implement the neural network inside the NXT, so that it would no longer rely on a PC for its operation. Several possibilities were available to implement the neural network inside of the NXT brick. One option would have been to try to fit Joone inside the NXT. Although this would have been the most versatile solution, it would have moved the focus away from an understanding of neural networks towards a more in depth knowledge of the Java language. Therefore the students decided to build their own neural network from scratch inside the NXT. This allowed them to gain a better understanding of the formulas that describe a neural network and an in-depth understanding of how to transform these formulas into Java code.

However, the NXT does not provide a user-friendly graphical user interface (GUI) that would enable users to easily manage the recorded audio samples and the training process. The students therefore decided to create the Neural Network Manager software (see Figure 3) for the PC that performs the training of the neural network.

Training the neural network on the NXT would in principle be possible as well, but of course at a much lower speed and only with an unfriendly user interface due to limitations of the NXT. It only has a small screen and four buttons to communicate with the user. A second reason for the preference of conducting the neural network

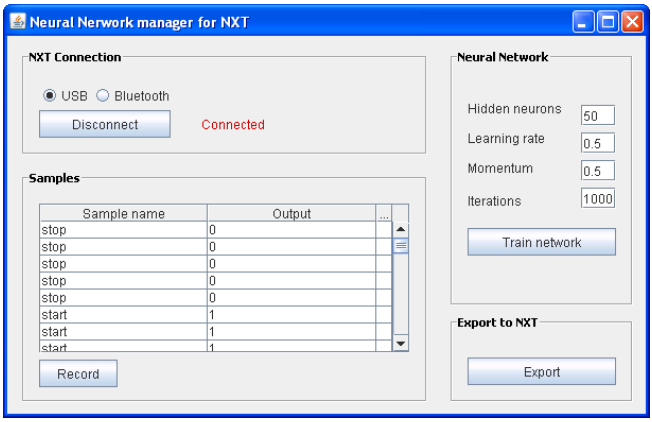


Fig. 3. Screenshot of the Neural Network Manager

training remotely on a PC, is that the network itself is unlikely to stand by itself. Most likely, it would be integrated into other software. This software needs to be created on the computer anyway and hence there was little reason to renounce the use of a computer. Once the neural network is trained, it can be transferred back to the NXT. It can then be used as a standalone application or as part in another program. The students were able to take advantage of their previous design education to create a highly usable GUI for the software. Hopefully, this will encourage other Lego users to take advantage of their software.

5.1 Neural Network Theory

We would like to conclude this case study by providing a short introduction to neural networks. This may allow readers that are not yet familiar with it to evaluate how much progress the students were able to make within three weeks.

Pattern recognition in general aims to classify data patterns extracted from raw data [1]. This is a very powerful tool to recognize classes of patterns where the raw data shows small variation or when the exact features are not known. In many cases using statistical information about the patterns or linear mathematical functions can do this. When both the data and the segmentation of the different patterns become more complex, neural networks are very suitable to perform pattern recognition tasks [11].

Neural networks in for example human brains consist of neurons connected through synapses forming a complex network. Artificial neural networks feature layers of neurons. A simple neural network at least has an input layer with neurons, an output layer with neurons and at least one hidden layer. All the neurons in a layer are interconnected through synapses to the next layer of neurons. Every neuron is connected to every neuron in the next layer (full synapses).

The synapses function as a weight factor and the neurons function as a mathematical function. Input can be fed into the neural network and is multiplied by the weight factors of the synapses. Neurons in the next layer apply a mathematical function, for example a sigmoid function to the sum of all the input values multiplied by their weight factor. This process repeats until the output neurons get a value. The output will return values that represent a specific pattern, at least when the weight factors are correct:

$$O_j = \frac{1}{1 + e^{-\sum_{i=1}^{m} x_i * w_{ij}}}$$

where,

O_j = output value of neuron j

m = neurons in previous layer

x_i = value of neuron i

w_{ij} = weight factor of the synapse between neuron i and neuron j

A common way to train a neural network is by means of backward propagation. Back propagation is a supervised learning method, which means that a set of input values coupled to desired output are used to train the network. The back propagation algorithm calculates the error signal by comparing the actual output with the desired output. It then uses the error signal to update the weights. The network is trained by repeating this iterative process until the actual output approximates the desired output:

$$\delta_j = \left(t_j - O_j\right) \cdot O_j \cdot \left(1 - O_j\right), \text{ where,}$$

δ_j = error signal for neuron j (in output layer)

t_j = desired output

O_j = actual output

$$\delta_j = O_j \left(1 - O_j\right) \sum_k \delta_k \cdot w_{kj}, \text{ where,}$$

δ_j = error signal for neuron j (in intermediate layer)

O_j = actual output

δ_k = error of (output) neuron k

w_{kj} = weight factor of synapse between neuron j and k

$$w_{ij}(t+1) = w_{ij}(t) + \eta \cdot \delta_j \cdot O_i, \text{ where,}$$

$w_{ij}(t+1)$ = new weight for synapse between neuron i and j (in all layers)

$w_{ij}(t)$ = current weight for synapse between neuron i and j

η = learning rate

δ_j = error signal on output

Oi = input signal of synapse

6 Conclusions

We described the embodied intelligence method to teach machine learning to design students. By using a tangible embodiment as a platform for machine learning, the environment of the machine-learning program does not need to be simulated. But more importantly, the embodiment provides the student with a tangible tool to understand and interact with a learning system. Lego Mindstorms NXT is a good platform for this embodiment. The Lego system allows the students to quickly build a machine and thereby enables students to focus on the machine learning. In addition Lego NXT provides a Java Virtual machine on which students can execute Java programs. Java is a widely used object-oriented programming language. The combination of the Lego construction system and Java is a very low hurdle that even students who do not have an affinity toward technology can overcome.

Many of the students played with Lego during their childhood. This positive memory might have lowered inner barriers that technophobic students might have built up. It might have allowed them to approach the course with a more open attitude and thereby increased the opportunity for learning. A second factor that might have had positive influence on the students is the behavior of the robots. The Crawler robot demonstrates that even simple learning behavior embodied in Lego has the power to create affection and empathy with human observers. This might have further motivated the students to experiment with the machine-learning program.

However, the embodied intelligence method does not only offer advantages for less technophile students, but it also offers enough room for advanced development. Within only three days certain students were able to build and use neural networks. They then continued to build their own neural network program from scratch, utilizing on the theory they learned in the preceding week. In the end, they were able to create neural network software that is user friendly enough for the general Lego enthusiast. As an example application, they built a voice command system, which enables the Lego NXT to operate as a stand-alone voice controlled device. Again, we have to emphasize that these were neither computer science students nor mathematicians. These were design students that normally create artifacts.

Only by enabling design students to understand, use and develop machine-learning systems, we can ensure that they will be able to create truly intelligent systems, products, and related services. The embodied intelligence teaching method can help achieving this goal and our experiences suggest that this has the potential to significantly help students who do not have an inhering affinity towards technology.

References

1. Voorhees, R.A.: Measuring what matters: competency-based learning models in higher education. Jossey-Bass, San Francisco (2001)
2. Weert, v.T.J.: ICT-rich and Competency Based Learning in Higher Education. In: Kallenberg, A.J., Ven, M.J.J.M. (eds.) The New Educational Benefits of ICT in Higher Education. Erasmus University Rotterdam, Rotterdam (2002)

3. Kyffin, S., Feijs, L.: The New Industrial Design Program and Faculty in Eindhoven. Designing Designers - International Convention of University Courses in Design, Milan (2003)
4. Bartneck, C., Jun, H.: Rapid Prototyping for Interactive Robots. In: Groen, F., Amato, N., Bonarini, A., Yoshida, E., Krose, B. (eds.) 8th Conference on Intelligent Autonomous Systems (IAS-8), pp. 136–145. IOS Press, Amsterdam (2004)
5. Bishop, C.M.: Pattern recognition and machine learning. Springer, New York (2006)
6. Nehmzow, U.: Mobile robotics: a practical introduction. Springer, London (2003)
7. Gasperi, M., Hurbain, P., Hurbain, I.: Extreme NXT: Extending the LEGO Mindstorms NXT to the Next Level. Apress, Berkeley (2007)
8. Solorzano, J.: LeJos (2002), `http://lejos.sourceforge.net/`
9. Sutton, R.S., Barto, A.G.: Reinforcement learning: an introduction. MIT Press, Cambridge (1998)
10. Watkins, C.: Learning from Delayed Rewards. King's College, vol. PhD. Cambridge University, Cambridge (1989)
11. Haykin, S.S.: Neural networks: a comprehensive foundation. Prentice Hall, Upper Saddle River (1999)

A Survey on Use of "New Perspective English Learning System" among University Students—Case Study on Jiangxi Normal University

Jing Zhang[1] and Min Li[2]

[1] School of Communication, Jiangxi Normal University, Nanchang, P.R. China
[2] WanNian Middle School, WanNian, P.R. China
zjing727@163.com

Abstract. As an e-learning system, New Perspective English Learning System is widely used in universities and colleges in China. However, little research has been carried on directed towards its effectiveness and correctness. This paper makes a case study on Jiangxi Normal University and adopts survey research as the predominant methodology and interview as secondary method. Via data collection and analysis, it examines respectively system effectiveness from respondents' satisfaction comments on five dimensions including students situation, learning resources, support from teachers, e-learning Support Platform and learning support service. Then it suggests an urgent need to solve the existing problems and improve the function of the whole NPELS Learning Hall platform in many ways. Finally some discussion and corresponding suggestions are proposed.

Keywords: New Perspective English Learning System, Survey, E-learning, Jiangxi Normal University.

1 Introduction

New Perspective English Learning System (NPELS) was designed and developed by Shanghai Foreign Language Education Press according to Instructional Requirements of College English Course (for trial implementation) published by Ministry of Education of PRC in 2003[1]. As an autonomous e-learning system based on the modern concept of teaching foreign language and supported fully by network technology, NPELS consists of three parts, i.e., interface of system administrator, interface of teacher management and Learning Hall for students. Among these, Learning Hall is the main part of the most significance because it is a public English learning platform open to the non-English major university students and it focuses on providing rich out-of-class resources of English learning and constructing an individualized autonomous e-learning environment.

So far as we know, over 180 China's universities and colleges have put this e-learning system into practice. However, there is no scientific evidence or data available which can indicate the effectiveness of NPELS use. Has it really improved

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 218–229, 2008.

autonomous English learning of university students? Does it serve as a vital complementary part for English classroom instruction as we always expect? To answer these questions in a scientific manner, a case study was made on the implementation of NPELS Learning Hall (hereinafter referred to as Learning Hall) in Jiangxi Normal University (JNU), which is a good case in point of having been using the system since 2006. We aim to examine the effectiveness and correctness of the design, development and operation of this e-learning system and share our experiences as a reference for other universities and colleges in China.

2 Theory

Although NPELS has been widely used in a number of China's universities, little research has been done directed towards its effectiveness. Up to now only three papers associated with NPELS can been searched in China National Knowledge Infrastructure Databases and these papers are conducted in a highly empirical way without full evidence from quantitative analysis. Moreover, only one master dissertation was found in Master Dissertations Full-text Database of JNU with its focus on NPELS-based autonomous learning strategy of university students instead of NPELS effectiveness evaluation. Hence a full-scale research conducted in a highly scientific manner must be undertaken. This research aims to provide theoretical and statistical analysis to examine different aspects of use of Learning Hall.

The theoretical constructs pertinent to this research are constructivism and the theory of web-based instruction evaluation.

Firstly, the constructivism believes that learners will construct their own knowledge and make a meaningful learning if they are put into a well-organized socially cultural-rich context [2].

Secondly, the theory of web-based instruction evaluation (He Kekang, 2002) is chosen to be the basis for this research because of its solid theoretical foundation and the fact that it has been proven successful in numerous empirical studies. From the perspective of this theory, the factors of web-based instruction evaluation include learners, teachers, web-based instructional Support Platform, instructional contents and learning support service [3].

3 Methodology

Learning Hall is characterized by its integrated platform components and typical student-learning process(see appendix 1). As the users of Learning Hall platform, Students can visit it and begin their learning after having finished a series of online pre-learning procedures involving login, different-level test and class choosing. Taking JNU as an instance, the student users of this university are required to enter Learning Hall for over ten hours accumulatively each semester, otherwise, they will fail to pass the written English exam of class-teaching. The state of using Learning Hall among the students of JNU and their satisfaction comments are the main contents of this study.

3.1 Hypotheses

Based on the theory of web-based instructional evaluation as the above discussed, and according to the practical situation of NPELS use, we assume five factors when examining the effectiveness and correctness of Learning Hall, which are respectively learners themselves, learning resources, support from teachers, e-learning Support Platform and learning support service. And we have the following hypotheses to explain what attributes a highly satisfactory e-learning system should have:

H1: Students are supposed to have authentic experience of entering Learning Hall and e-learning before they make any comment towards it.

H2: The effectiveness of learning resources is positively associated with learning content providing, structure and guiding of the interface and interaction function.

H3: The satisfaction to support from teachers is positively associated with their attitude and how they offer materials to students.

H4: The effectiveness of e-learning Support Platform is positively associated with its instructional and technical function.

H5: The satisfaction to learning support service is positively associated with "help" function.

3.2 Method and Objects

Survey research is adopted as the predominant methodology in this research. Based on the above five hypotheses and related theories, a questionnaire was designed includes five basic dimensions—students situation, learning resources, support from teachers, e-learning Support Platform and learning support service, of which detailed items were put forward as sub-dimensions. After a trial small-scaled distribution of questionnaire draft and further modification and finalization of the questionnaire design, random sampling was used when circulating the final questionnaires face-to-face. Via Questionnaire-investigation, as the section four shows, scientific data were collected and analysis was done.

As Learning Hall users, the non-English major freshmen and sophomores in JNU are the investigated objects (Generally speaking, only freshmen and sophomores in China's universities are required to attend English class at school except for the students majoring in English or other foreign languages). Totally 180 questionnaires were distributed and 163 of them were collected with questionnaire return rate of 90.55%, among which 151 questionnaires were valid with the effective rate up to 92.64%. All the 151 questionnaires were collected and analyzed. In addition, some of the respondents were interviewed as a complementary research method so that a comprehensive result can be obtained and an informed judgment can be made.

4 Statistical and Data Analysis

4.1 Students Situation

First of all, through collecting basic information of the students, the respondents cover a wide range of different majors background. As Table 1 shows, the rates of the

investigated students are in a fairly balance seen from their grade and sex. The rate of the art and science students, however, shows a visible difference, which is mostly because the enrollment of science major students surpasses considerably that of art major students in JNU recent years. Hence the validity to this research results is ensured.

Table 1. Investigated students' background

Classification	Grade		Sex		Major		
	Freshmen	**Sophomores**	**Boys**	**Girls**	**Art**	**Science**	**Others**
Number	72	79	71	80	44	102	5
Rate(%)	47.7	52.3	47.1	52.9	29.1	67.6	3.3

The other single question asking when you usually login Learning Hall is used to measure the students' attitude towards learning in Learning Hall. We find out that 39.87% of the respondents login in normal times whereas 28.76% login typically when an English exam is approaching, left the rest 31.37% login in uncertain time. It shows that nearly one third of the students are motivated to use Learning Hall by exams rather than their interest.

4.2 Learning Resource

Learning content. The learning process in Learning Hall can be divided into several stages as follows—login, different-level test, class choosing, modular learning, modular test, promotion to upper module and promotion to upper grade.

Firstly, seen from the data about "the most satisfactory stage and the most dissatisfactory stage in the whole learning process", it is clear that the students are most satisfied with "modular learning" and most dissatisfied with "modular testing"(see Figure 1). Meanwhile it is found easily that some of the students don't express an explicit opinion by choosing "none", causing the rate of most satisfactory reaches 29.14% and that of most dissatisfactory 24.50%, which produces a close result.

When being asked the reason of choosing "none", some students explain that they are not familiar with the whole learning process in Learning Hall. Some of them only enter and keep opening the platform interface for a long time without further visit to the learning content patiently and earnestly so that they can meet the demand from the Academic Affair Office of JNU (i.e. ten hours' log on Learning Hall accumulatively each semester) and avoid failure in written English exam of class-teaching. Then why the most dissatisfactory stage is modular test? In the interview, the students give a common explanation that should be taken into account. Learning Hall requires that the students can come into "modular test" stage only after having learned in the prior "modular learning" stage for a fairly long time. Hence it keeps many students out of reach of "modular test" stage, leading to a dissatisfied inclination.

When being asked the reason of choosing "none", some students explain that they are not familiar with the whole learning process in Learning Hall. Some of them only

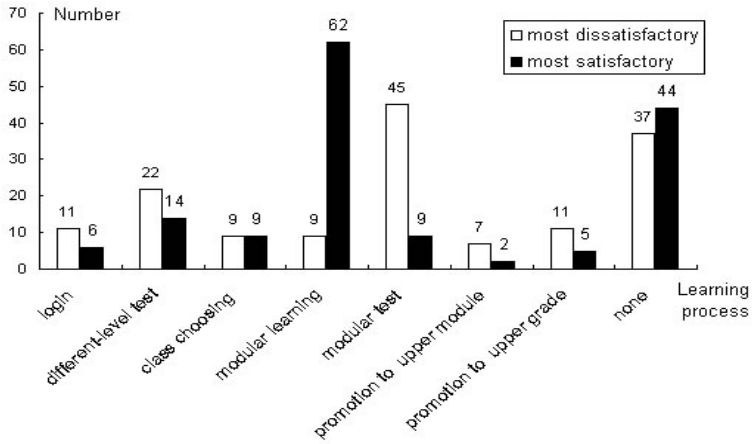


Fig. 1. Most satisfactory stage and most dissatisfactory stage in the whole learning process

enter and keep opening the platform interface for a long time without further visit to the learning content patiently and earnestly so that they can meet the demand from the Academic Affair Office of JNU (i.e. ten hours' log on Learning Hall accumulatively each semester) and avoid failure in written English exam of class-teaching. Then why the most dissatisfactory stage is modular test? In the interview, the students give a common explanation that should be taken into account. Learning Hall requires that the students can come into "modular test" stage only after having learned in the prior "modular learning" stage for a fairly long time. Hence it keeps many students out of reach of "modular test" stage, leading to a dissatisfied inclination.

"Modular learning" is the most important stage in the whole process and it contains various units. The question of "the most satisfactory unit and the most dissatisfactory unit in modular learning" is asked and data is analyzed (see Figure 2).

Figure 2 displays that totally the students are most satisfied with "supplementary resources for learning" with a high rate of 43.70%. It has some relevance into the fact that the students have easy access to all kinds of rich complementary resources for learning in this stage. And the most dissatisfactory unit is "meeting teacher". Ideally, by "meeting teachers", students can make an online appointment with a desired teacher first and then get a face-to-face instruction from the teacher. But the premise of fulfilling this is that students must pass unit test prior to meeting teachers. However, virtually few students meet teachers due to the difficulty in passing unit test for various reasons although its function is quiet desired by students.

Structure and guide of Learning Hall. Among the 151 respondents, 79 students (accounting for 52.32%)think the structure and guide of Learning Hall is designed rationally which always prevent them from getting lost in the process of e-learning; 39 students (accounting for 25.83%) think it unreasonable and the rest 25 students (accounting for 21.85%) are not clear about this point.

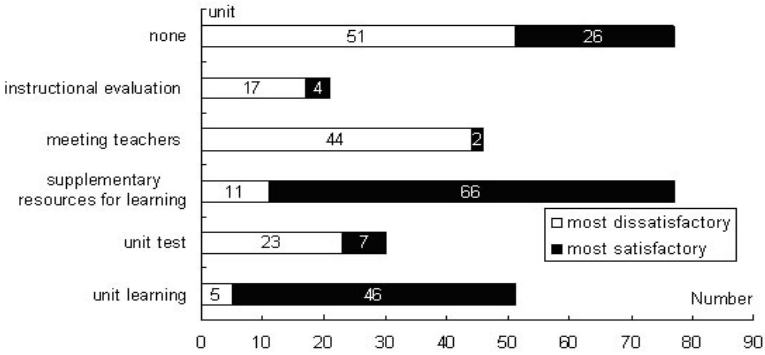


Fig. 2. Most satisfactory unit and most dissatisfactory unit in modular learning

The data reflects that totally the structure and guide design of Learning Hall is well accepted and effective. However, "getting lost in information and network" will also occur when learning resources are presented unreasonably. Therefore structure and guide design and learning resources design both play a key role in preventing getting lost and producing effective learning result.

Interaction. Interaction is a significant sub-dimension of learning resources and it is embodied in the following questions and data analysis. As Table 2 shows, nearly 60% students approve that Learning Hall provides a fully autonomous learning setting but at the same time almost half students feel somewhat lonely during learning in Learning Hall. What's more, over 70% students are positively interested in trying to use "meeting teachers" if it is available.

Table 2. Interaction state when learning in Learning Hall

Interaction	Yes (%)	No (%)	Not clear (%)
Do you always feel lonely when learning in Learning Hall?	46.00	35.33	18.67
Do you feel a fully autonomous learning setting when learning in Learning Hall?	59.60	23.84	16.56
Are you interested in having a try on utilizing "meeting teachers" function if it is open to you?	75.50	12.58	11.92

4.3 Support from Teachers

Teachers' attitude. In e-learning environment, a teacher plays a different role in comparison to that in a traditional classroom-teaching setting. A teacher becomes a supporter and facilitator of students' e-learning instead of a transmitter of information because e-learning platform can do most work of passing on to students all kinds of

information and knowledge [4]. In the process of autonomous learning, a teacher is supposed to facilitate students in developing a couple of capabilities such as setting learning goal, choosing learning content and strategy, arranging schedule, self-regulating learning process and self-assessing learning effect, etc [5]. Hence a teacher's attitude towards students may be embodied in how he or she is concerned about students, whether he or she organizes interactive communication among students effectively and whether he or she gives a timely feedback to their performance. To identify the teachers' attitude, the investigated students give their response as Figure 3 shows.

Unoptimistically, a relative large proportion of investigated students (55.62%) think their teachers seldom show concern for their e-learning in Learning Hall, which is 14 times of the number of the students (3.97%) who feel frequent concern from the teachers! Clearly the English teachers involve themselves in the interaction with their students in a low extent. And the teachers and the school cannot ignore this result.

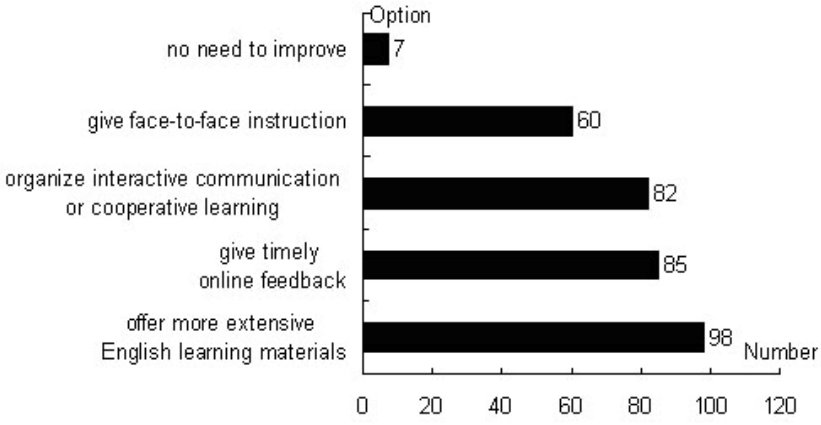


Fig. 3. On which aspects should teachers improve their function in Learning Hall

As for the Multi-optional question of "on which aspects should teachers improve their function in Learning Hall", Figure 3 displays that an overwhelming majority among students hopes their teachers to improve their function on various aspects. The top three aspects they want teachers to improve are respectively to offer more extensive English learning materials, to give timely online feedback and to organize online interactive communication or arrange cooperative learning activities among students.

Offering learning materials. The second dimension presenting supports from teachers is the situation of their offer learning materials to students, which are mainly embodied in the quantity and the update frequency of the uploaded resources by teachers. Figure 4 below reflects that the investigated students are not very satisfied with the uploaded materials from their teachers. It's general that the uploaded materials are not in variety and lacks frequent update.

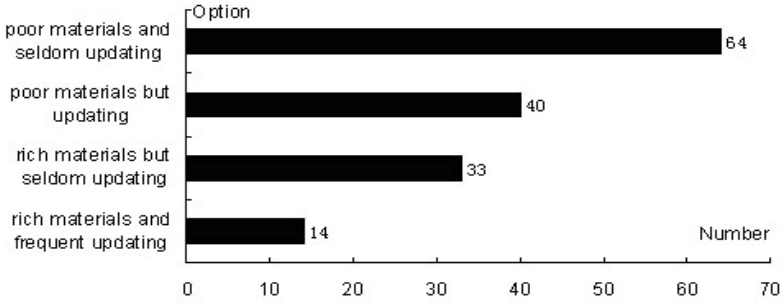


Fig. 4. Situation of teachers' uploading learning material

4.4 E-Learning Support Platform

As the forth dimension of the questionnaire, E-learning support platform's function can be evaluated instructionally and technically.

Instructional function. Instructional function of e-learning support platform is typically reflected in its support to students' learning by providing various resources and tools. To the question of "whether the resources and tools of Learning Hall can meet your needs of autonomous learning", 52.32% students choose "no" whereas only 22.51% students choose "yes", left the rest 25.17% have an unclear opinion on it. In the further interview, some students complain that many tools perform practically no function like forum, meeting teachers, not to mention tools for exploratory or collaborative learning.

Technological system. When being asked " how often tech breakdown happens when you are using Learning Hall", the number of respondents who choose "often" or "occasionally" accounts for 97% in comparison to the 3% students who says they seldom meet tech breakdown. And the top three kinds of tech breakdown told in the interview are failure to log on, low speed connection to Internet and getting disconnected with line suddenly.

4.5 Learning Support Service

Learning support service is as important as the curriculum and media. It plays an integral role in teaching and learning although it influences e-learning result in an indirect way.

Learning support service involves typically providing various helps or instructions for students in the process of e-learning. However up to now Learning Hall just provides "system using help". A single question asking respondents whether they are satisfied with the help in Learning Hall is used to measure the learning support service. Among the investigated 151 students, 34 students (22.52%) think the help service isn't helpful and 112 students (74.17%) think the help service is helpful but not enough because some

necessary helps (e.g. English learning strategic helps and e-learning helper) is not included, left only 3.31% respondents' affirmative opinion on the help service.

5 Research Result and Suggestion

Through the above data collection and analysis, undoubtedly NPELS Learning Hall is faced with many various problems in both its initial functional design and practical implementation in school. There is an urgent need to solve these problems and improve the function of the whole Learning Hall platform in many ways. Hereby some discussion and corresponding suggestions are proposed.

5.1 Develop an Effective Learning Process Monitoring and Assessment

Both the questionnaire and interview reflect that some students show a very ambiguous opinion on many basic attributes of Learning Hall such as learning process, internal structure and function by choosing the answers like "not clear", "None" or "I don't know". It is mostly because it's too easy for the students to meet the demand from the school—if a student has stayed online in Learning Hall for over ten hours in one semester accumulatively and then he or she will be qualified whatever his or her attitude and performance in the whole learning process. We believe only a time demand is not enough. An effective learning process monitoring and assessment is neglected in both the design and the implementation of Learning Hall and it is urgent to be developed at the earliest.

5.2 Create a Desirable E-Learning Environment

As the above mentioned, from a perspective of a constructivist, learners will construct their own knowledge and make a meaningful learning if they are put into a well-organized socially cultural-rich context. And this context may be more suitable to the students who practice autonomous e-learning in NPELS. However, NPELS does much work on providing learning resources rather than design a desirable learning environment, to which point the students show their complaint in both questionnaire and interview.

Therefore it is necessary to construct a desirable e-learning environment as soon as possible. Here some suggestions are proposed.

Despite that the reasonable structure and guide of the interface is approved by students due to its function in avoiding getting lost in Learning Hall, the interface may be designed more friendly and creatively so as to be more attractive to students. "For example," a student said in the interview, "the interface can be designed similarly in the style of OICQ (a native chatting software) interface. I'll be much more interested if I can see how many students there are in Learning Hall simultaneously and identify who are they." As the student said, it's no wonder e-games and e-chatting are so popular in youngsters because of their friendly and creative interface.

On the other hand, generally the interaction state is not very optimistic. Referring to the theory of three types of interaction (Michael G·Moore, 1989), learner-content

interaction of Learning Hall should be more designed and developed. Learner-learner and learner-instructor interaction is also a prerequisite and need to be improved at the earliest if an effective and active learning result is expected [6]. Hence meaningful learning topics or cooperative learning activities should be produced in virtual community and interactive forum, in which way students can practice oral English in a multiple English chatting cyberspace.

5.3 Improve Supports from Teachers

To some extent the teachers in NPELS should be accountable to the inactive performance of the students. The teachers should more care for students and analyze what the students need by participating in their learning activities and sharing their learning experiences. Rich materials uploading and frequent updating should be made. Moreover, to arrange cooperative learning activities among students, and to give timely online feedback should also be indispensable. In this way, we believe students will view Learning Hall as an effective approach to improve English learning.

5.4 Optimize E-Learning Platform Support and Learning Support Service

Some problems on NPELS itself are discovered in the survey and interview which the platform developers should take into account. On the other hand, e-learning platform urgently need to optimize its function in providing interactive and cooperative learning tools such as forum, Netmeeting and chatting room, etc. They should be included and work practically instead of being an empty shell in NPELS. Particularly "meeting teachers" function should be more accessible and open to students without the constraint of passing a test. On the other hand, learning support service should be more individualized and works in humanity manner. Besides the system using help, e-learning strategy helper, English learning method assistant should be embraced in the whole learning support service. In this way, students will be motivated to make an active learning.

6 Conclusion

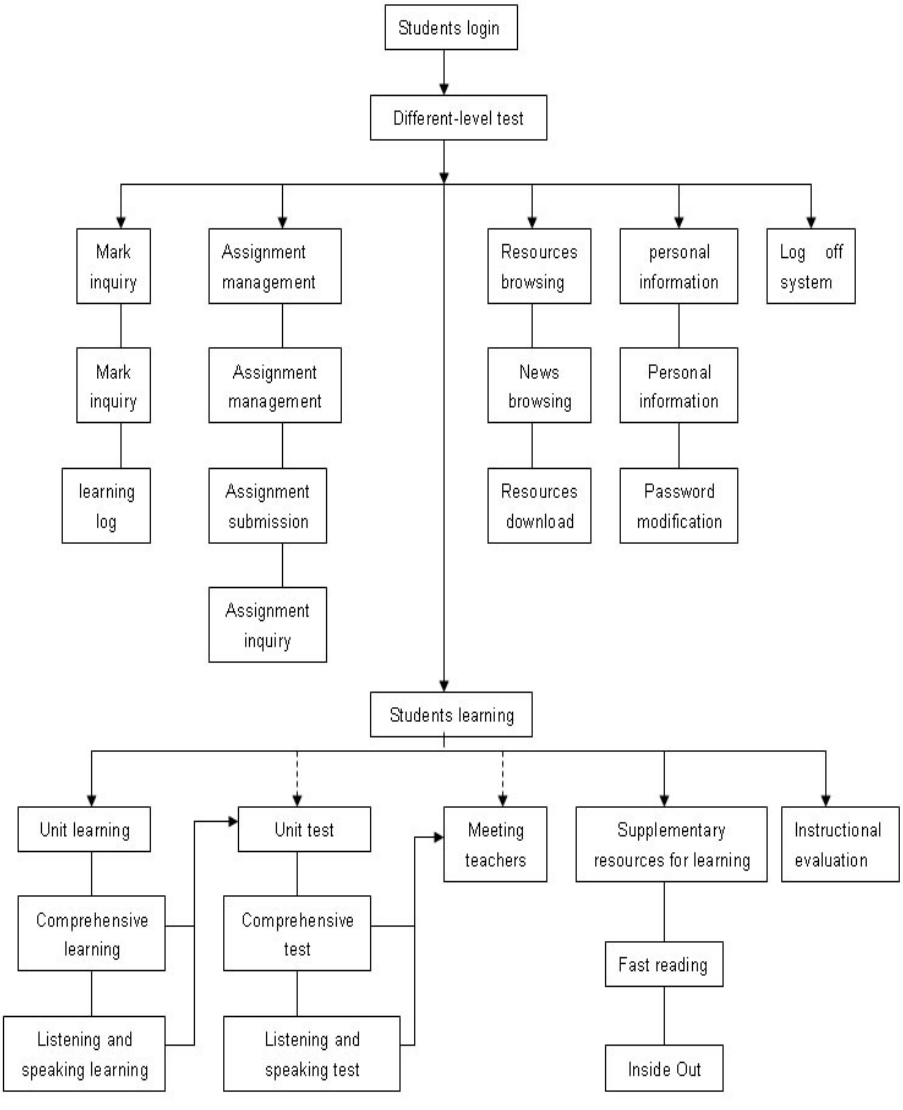
This research, focusing on the case study of JNU, reflects the common problems sharing among the universities and colleges of the same kind in China. This paper suggest a fully improvement be obtained on the aspects of system management, teachers' support, platform function and learning support service. Hopefully, the effectiveness and correctness of NPELS Learning Hall will be promoted after a series of improvement and optimization in overall function of the system.

Like most research, this study is not without limitations. The results in this study may have limited to the universities of different types in China. Future research could choose various kinds of universities or colleges as samples so that a more comprehensive result may be achieved.

References

1. Shanghai Foreign Language Education Press: NPELS End-user Manual (Student)(January 30, 2007), http://jwc.jxnu.edu.cn/NPELS/User/Login.aspx
2. Wertsch, J.V., Toma., C.: Discourse and learning in the classroom: A Socialcultural approach. In: Steffe, L.P., Gale, J. (eds.) Constructivism in education, pp. 159–174. Lawrence Erlbaum Associates Publishers, Mahwah (1995)
3. Kekang, H., Wenguang, L.: Education Technology, p. 363. Beijing Normal University Press, Beijing (2002)
4. Xinmin, S.: Learning Sciences & Learning Technologies, pp. 102–103. Higher education Press, Beijing (2004)
5. Shengquan, Y.: Web-based instruction evaluation model construction (March 20, 2007), http://www.qdedu.gov.cn/jiaoyuguanli/lwjc/article/Article3147.htm
6. Moore, M.: Three Types of Interaction (March 19, 2007), http://cdl.panam.edu/dayoung/professional/moorethreetypesofinteraction.htm

Appendix 1: NPELS Learning Hall Platform Structure



Evolving Game NPCs Based on Concurrent Evolutionary Neural Networks

XiangHua Jin, DongHeon Jang, and TaeYong Kim

Department of Image Engineering, Graduate School of Advanced Imaging Science Multimedia, and Film, Chung-Ang Univ., 221 Heuseok-Dong Dongjak-gu, 156-756 Seoul, South Korea
hyanghwa_kim@naver.com, tellamon@gmail.com, kimty@cau.ac.kr

Abstract. Evolutionary Artificial Neural Networks (EANNs) has been highly effective in Artificial Intelligence (AI) and in training Non-Player-Characters (NPCs) in video games. An important question in training NPCs in games is how we can choose the appropriate way to make NPCs smart. We focus on (1) choosing a principled method of high dimensional data space, (2) designing adaptive fitness functions which can make the proper evolution. In this work, we describe the Concurrent Evolutionary Neural Networks (CENNs) based on EANNs for competitive team game playing behaviors by teams of virtual football game players. We choose Darwin Platform as our test bed to show its efficiency. The Red team and the Blue team are competing in the soccer field, the field players in Red team are evolved during the virtual game playing. The experimental results show that the Blue team programmed by Rule-Based System leads the evolution successful.

Keywords: Evolutionary Artificial Neural Networks, football game NPCs.

1 Introduction

Game Industry, as a complex of medium which includes story telling, artworks, sound and techniques such as Math, Physics, rendering, has been developed fast enough to be the mainstream of cultural industry. When the Graphics and Sound techniques have reached at certain level, the game users want the games more realistic and interesting. For these reasons, the Artificial Intelligence (AI) technique has become more and more important in game industry [1]. As computer games become more complex and consumers demand more sophisticated computer-controlled NPCs, the game developers are required to place a great emphasis on the AI. In order to make the game more realistic and smarter, high level AI has been strongly demanded by game users since 1990s. For that, game programmers can educate/train game NPCs using various AI techniques. When designing the specific AI, it is important to choose adaptive AI method/algorithm. If we use low-level algorithm, people always can predict the action of the NPCs. It will make the game boring. Recently there has been much interest in combining evolutionary algorithms and artificial neural networks [2], [3], [4], [5], [6], [7]. Evolutionary Artificial Neural Networks (EANNs) has been highly effective in training NPCs in video games, because EANNs can evolve the whole structure of the Neural Networks which can be represented as behavior of NPCs [16].

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 230–239, 2008.

As we can know from the real football games, every player not only has to consider his own position, his team members' positions, enemy team members' positions, and ball position and so on, but also has to decide his action to win the game. When we simulate the whole things in digital football games, it is too hard to check every circumstance. In other words, we hardly use rule-based systems or Finite State Machine (FSM) to design AI. However, it is possible to use EANNs. We can encode the virtual player in a form of Neural Network controller, all the positions we described above set to be neural inputs and the actions to be the outputs.

As we described above, the player must pay attention to so many positions that the evolutionary algorithm can be considered as the optimization problem of the high dimensional search space. We focus on (1) choosing a principled method of high dimensional data space, (2) designing adaptive fitness functions which can make the right evolution. We describe the Concurrent Evolutionary Neural Networks (CENNs) based on EANNs for competitive team game playing behaviors by teams of virtual football game players. We choose Darwin Platform as our test bed to show its efficiency. Darwin is ergonomics A.I. game platform developed in the project of Ergonomics Game Intelligence [8].

2 Related Works

2.1 EANNs Used in Training Game Agents

In the area of using EANNs to design game AI, one of the main problem is finding suitable structures for evolving adaptive game NPCs [16].

Enforced Sub-Populations (ESP) [9], is a kind of EANNs system based on Symbiotic, Adaptive Neuro-Evolution (SANE) [10]. In ESP, populations of neurons are evolved to form a neural network. It differs from other EANNs systems in that it evolves a population of neurons instead of complete networks. A neuron is selected from each population to form the hidden layer of the network, which is then evaluated on the problem and its fitness passed back to the participating neurons. These neurons are combined to form neural networks that are then evaluated on a given problem. ESP differs from SANE in that neurons for the different positions in the network are evolved in separate subpopulations. ESP has been used in training the agents in some video games [11] and was shown to be significantly faster than other EANNs methods.

Another EANNs method used in developing game AI is NeuroEvolution of Augmenting Topologies (NEAT) [15]. NEAT starts from the neural networks with small structures and becomes increasingly sophisticated over generations. This technique is appropriate for the real-time game where the agents are repeatedly created and eliminated.

In sum, the approach in using EANNs to design game AI is a useful integration. In our work, we try to design the AI of football game NPCs where team-cooperative behavior is expected and propose a method which can evolve the weights while decreasing the number of connection of the neural networks, as we will discuss in Section 3.

2.2 Darwin

Darwin, a game platform, based on agent that developers embody AI easily and capable of proposing AI test with module that makes them find strategic action and

then evaluate achievement results through making agent used strategic module that Darwin offers.

Darwin consists of AI-Level and Game-Level (Fig. 1 shows the framework of Darwin Platform). Game designer can make DLL using the template that Darwin provides. When AI designer develops a DLL, Darwin-Manager in Game Level provides the main program functions of Neural Networks, Genetic Algorithms, FSM, basic EANNs and so on. DLL developed by game AI designer can be loaded by Darwin-Loader in Darwin-Manager and then the NPCs' action can be seen from Darwin-Viewer.

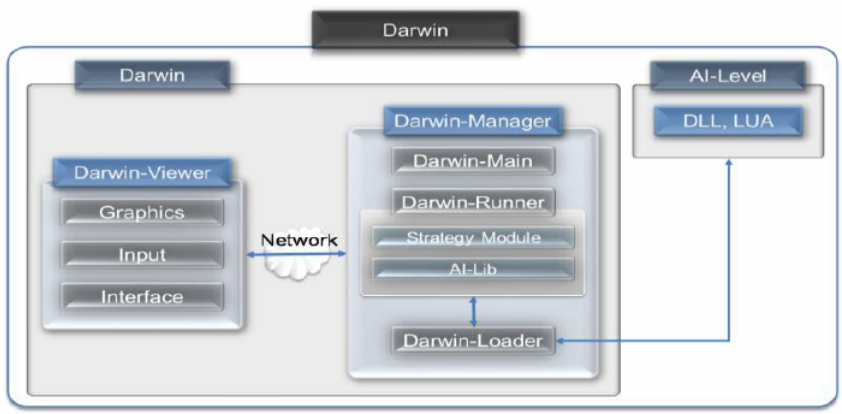


Fig. 1. The framework of Darwin Platform

3 Concurrent Evolutionary Neural Networks

Concurrent Evolutionary Neural Networks (CENNs) deletes the connection of the neural networks because there may be some poorly performed ones from neurons to neurons [7]. To reduce the search space, we delete the connections which perform worst in games during the weights' evolution.

In the initial step, populations of full-connection neural networks are formed by random weights. In our work, one neural network contains 18 input neurons, 5 output neurons and 8 hidden neurons. The number of the weights in one neural networks is (18+6+1)*8 = 200. If we want to evolve the structure with full-connection in our experiment, the length of one gene will be 200. In other words, a gene is a set of 200 numbers that may lead the search space too large to evolve. In order to solve the problem, we suggest CENNs and we choose Real-Number Representation [7] as our encoding method (Figure 2).

The neural networks are then evaluated according to take part in the gaming domain. The fitness value is passed back to the participating weights. The main process is as described below.

1. **Initialization.** Create the initial population of neural networks with the random weights.

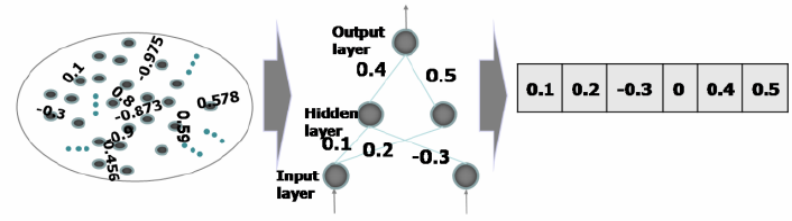


Fig. 2. Real-Number Representation Encoding

2. **Evaluation of each network.** Evaluate each neural network by playing the game. At the end of each game, the fitness function evaluates the AI player based on goal scoring and contribution factor and so on.
3. **Temporal elite selection.** Choose the best *n* neural networks to participate in the next connection evaluation step.
4. **Evaluation of each connection.** Re-evaluate the selected elite network whose one of input-hidden connections is disabled. Disabling of connection occur consecutively for all input-hidden connections. For each disabled connection, average fitness score for *n* neural networks is evaluated.
5. **Deletion of selected connection.** Whether delete the connection or not is measured by the method based on the formula of estimating node relevance within a trained network [14]. It consists of evaluating the effect that removing the node has over the error. In our work, the tested connections with lower fitness score in step 4 are selected and deleted if its fitness score is smaller than threshold error. In each generation, we delete two connections.
6. **Evolution step.** In this step, we evolve the weights of neural networks. To optimize weighting values, we use Simplified Differential operator (SADE) [5], [12] instead of crossover. In such situation, it is not appropriate to use crossover operator because of its high dimensional data space. SADE operator adopts Real-coded Genetic Algorithm and the main operators we use in our experiment in this process are as described below. By these operators, the number of the neural networks will be doubled.
   1) Simplified differential operator: Let $CH_i(t)$ be the *ith* chromosome in a generation *t*,
   $$CH_i(t) = (ch_{i1}(t), ch_{i1}(t), \cdots, ch_{i1}(t)) \tag{1}$$
   where, *n* is the number of variables of the fitness function. Then the simplified differential operator can be written as:
   $$ch_{ij}(t+1) = ch_{pj}(t) + CR(ch_{qj}(t) - ch_{rj}(t)) \tag{2}$$
   where ch_{pj}, ch_{qj} and ch_{rj} are the *jth* coordinates of three randomly chosen chromozomes and *CR* is so-called cross-rate.
   2) Local mutation: if a certain chromozome was chosen to be locally mutated, all of its coordinates are altered by a random value from a given range.
7. **Selection.** Randomly choose two networks and then reject the worse one, the generation count is increased by one.

8. **Check criterion.** Repeat 3-7 until the fitness value reaches at the criterion.
9. **Save and load.** Save the best architecture in the whole process and load in the game.

4 Experiment

4.1 Experimental Discipline

In the virtual soccer game, played by Blue team and Red team, the two field players in Red team will be evolved during the game play. The Blue team is pre-programmed by Rule-Based System to help the neural network players evolve successful. We name the target field players as NEPlayer evolving with CENNs algorithm.

NEPlayer has two separate neural network systems depending on the status that the player has the ball or has not the ball.

1) When NEPlayer has not the ball: The network takes the inputs from the coordinates of friendly players, enemy players and the ball. Darwin test bed provides the sensing ability of each positions of the players or the ball. The sensored inputs are transferred to the outputs through the weighted nodes of the hidden and output neurons resulting from sigmoid functions. The outputs control the moving direction or target position of NEPlayer without the ball. The combination of output and the corresponding directions is shown in Table 1.

Table 1. The combination of neural outputs and its behavior

Output	Behavior
000	Move to the midpoint between the ball and the enemy goal
001	Move to the midpoint between the nearest friendly player and ball
010	Move to the midpoint between the second nearest friendly player and ball
011	Move to the midpoint between the third nearest friendly player and ball
100	Move to the ball position
101	Move by 0 degree direction
110	Move by 120 degree direction
111	Move by 240 degree direction

2) When NEPlayer has the ball: The network takes the same inputs as in the previous case except the ball position. The outputs control not only the moving direction or target position of NEPlayer but also the power to kick the ball. We have prepared different expected behavior from the Table 1. For example, if the first to third outputs are 000 and the fourth one and fifth one is 00 (the fourth, fifth ones are one of 00~11 controls the kicking power of the ball), then it will kick the ball at specified direction in Table 2 with 1/4 multiplied by the maximum kick power. The combination of output and the corresponding directions is shown in Table 2.

4.2 Design of Fitness Function

It is also very important to design the adaptive fitness function in the evolutionary processing. In our experiment, we want the NEPlayer not only win the game but also

Table 2. The combination of neural outputs and its behavior

Output	Behavior
000	Kick to the nearest friendly player
001	Kick to the second nearest friendly player
010	Kick to the third nearest friendly player
011	Kick to the midpoint between the third nearest enemy player and the farthest enemy player
100	Kick to the midpoint between the second nearest enemy player and the farthest enemy player
101	Kick by 0 degree direction
110	Kick by 120 degree direction
111	Kick by 240 degree direction

act as a human player. We must consider as many situations as possible that can be taken place during the game play. Like the Neural Network architecture, the fitness function has two separate cases depending on the status that the player has the ball or has not the ball.

The game is initialized when either team get a goal. At the initialization step, the fitness functions are evaluated considering the three situations as follows.

1) When NEPlayer has not the ball: In this case, we can calculate the fitness based on which team has the ball and how well the members are distributed. The all distances between each player are summed and normalized to calculate the distribution bonus in Darwin Platform. Let f be the fitness score, N_F be the number of the friendly team has the ball until any player gets a goal, N_E be the number of the enemy team player has the ball until any player gets a goal, N_{Bonus} be the number of the players distributed well, T be the frames from starting to the moment of getting a goal, T_F be the time limit. In our experiment, considering that it takes about 1800 frame to get a goal in the Darwin Platform, we set the time limit by 3000 frame. The formulas are shown below:

i. When friendly team gets a goal:

$$f = (N_F - N_E + N_{Bonus})/T \tag{3}$$

ii. When enemy team gets a goal:

$$f = -(N_F - N_E + N_{Bonus})/T \tag{4}$$

iii. When neither team gets a goal until time limit:

$$f = 0.5 * (N_F - N_E + N_{Bonus})/T_F \tag{5}$$

2) When NEPlayer has the ball: In this case, the fitness function evaluates its fitness according to the following rules. If the teammate gets the ball, it adds 2 points to the fitness value but if the enemy player gets the ball, it loses 2 points. When nobody gets the ball, it calculate the nearest player to the ball then, it will add or subtract 1 point depending on the nearest player's team. When the game ends, these values are summed up and set to the fitness value of each game. Let f be the fitness score, N_F be the number

of the NEPlayer correctly kick to the friendly team member until any player gets a goal, N_E be the number of the NEPlayer kick to the enemy team player until any player gets a goal, N_{NF} be the number that the nearest player to the ball is friendly team player when nobody takes the ball, N_{NE} be the number that the nearest player to the ball is enemy team player when nobody takes the ball. The formulas are as follows.

i. When friendly team gets a goal:

$$f = (S_F * N_F + S_E * N_E + S_{NF} * N_{NF} + S_{NE} * N_{NE}) / T \tag{6}$$

ii. When enemy team gets a goal:

$$f = -(S_F * N_F + S_E * N_E + S_{NF} * N_{NF} + S_{NE} * N_{NE}) / T \tag{7}$$

iii. There is no team get a goal during time limit:

$$f = 0.5 * (S_F * N_F + S_E * N_E + S_{NF} * N_{NF} + S_{NE} * N_{NE}) / T \tag{8}$$

All the parameters used in our experiment is as shown in Table 3.

Table 3. Parameters use in the experiment

Parameter Type	Value
Initial value of each weight	-1~1
Number of the initial Neural networks	1280
CR	0.2
MR	0.5
The local Mutation Range	0.25

Table 4. Approximate averaging number of the evaluation of CENNs, Binary Genetic Algorithms (Binary GAs) and ESP

Algorithm	Generations	Evaluations
Binary GAs	500	1280,000
ESP	250	640,000
CENNs	100	299,200

4.3 Experimental Results

In order to evaluate the evolution process and the result, we evolve the same structure of Neural Networks using Binary Genetic Algorithms [13], ESP and CENNs. The approximate number of the averaging result is as shown in Table 4.

Table 4 shows that CENNs is an appropriate system in evolving football game NPCs. CENNs takes much fewer evaluations than Binary GAs and ESP, showing that in the high dimensional optimization problems sometimes we need concurrent events to make the evolution faster.

To prove its usability and efficiency, we have matched the untrained NEPlayers and the trained NEPlayer against the pre-programmed players with FSM for 100 games. The match result of untrained NEPlayers against the FSM Players is as shown in figure3.

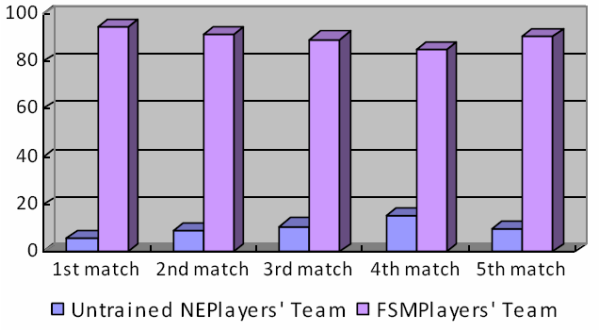


Fig. 3. The match result of untrained NEPlayers against the FSM Players

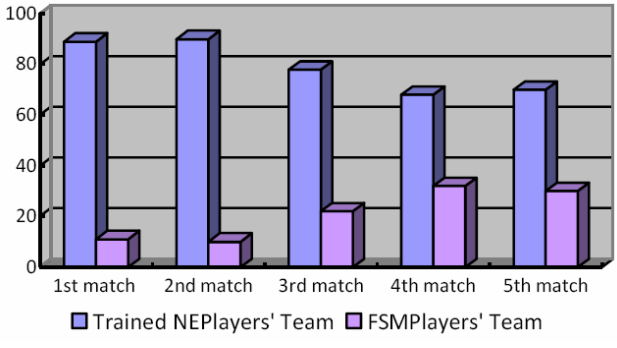


Fig. 4. The match result of trained NEPlayers against the FSM Players

Darwin Platform provides save and load NE Players which get the best fitness score in the game (Fig. 5 shows the screenshot). The match result of trained NEPlayers with the best fitness value against the FSM Players is as shown in Fig. 4.

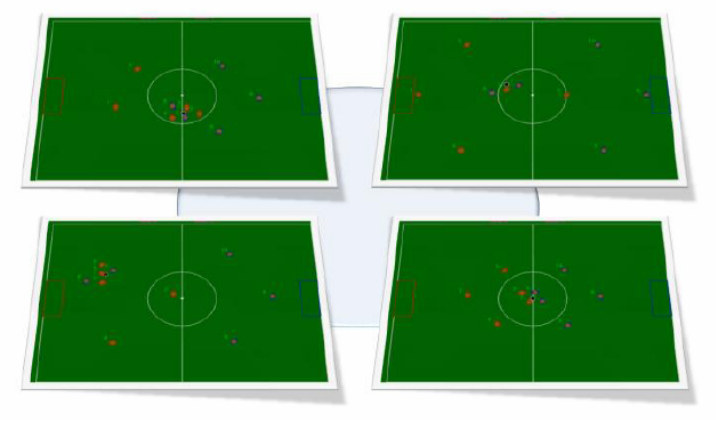


Fig. 5. The screenshot shows the NEPlayers play in Darwin Platform

As we can know from all the results above, CENNs not only can speed up the evolutionary computation time, but also is an efficient algorithm with the adaptive fitness function in video football game. Although enemy team becomes stronger, the NE-Players trained by CENNs act cooperatively in team plays and win the games.

5 Discussion and Future Work

As shown in the experiment, CENNs algorithm is robust solution for evolving football game agents. Although the number of weights in neural network is more than 100, the test on Darwin Platform shows satisfied result. For wide range of game applications, the number of weights must be carefully considered and the gaming situation can be more complex. Larger number of weights causes the larger search space which may lead to poor performance; more complex situation may causes more troubles of designing fitness function. The future work is focusing on the proper designing of fitness function and algorithm to speed up in real time gaming environment.

Acknowledgments

This research was supported by the ITRC (Information Technology Research Center, MIC) program and Seoul R&BD program, Korea.

References

1. Rollings, A., Morris, D.: Game Architecture and Design (2000)
2. Kitano, H.: Designing neural networks using genetic algorithms with graph generation system. Complex Systems 4, 461–476 (1990)
3. Koza, J.R., Rice, J.P.: Genetic generalization of both the weights and architecture for a neural network. In: International Joint Conference on Neural Networks, vol. 2, pp. 397–404. IEEE, New York (2000)
4. Liu, Y., Yao, X.: A population-based learning algorithm which learns both architectures and weights of neural networks. Chinese Journal of Advanced Software Research 3(1) (1996)
5. Hrstka O, K.: A Search for optimization method on multidimensional real domains. In: Contributions to Mechanics of Materials and Structures. CTU Reports, vol. 4, pp. 87–104. Czech Technical University, Prague (2000)
6. Whitley, D., Starkweather, T., Bogart, C.: Genetic algorithms and neural networks: Optimizing connections and connectivity. Parallel Computing 14, 347–361 (1990)
7. Yao, X., Liu, Y.: Evolving artificial neural networks. Proceedings of the IEEE 87(9) (September 1999)
8. Im, C.-S., Kim, T.Y., Um, S.-W., Baek, S.H.: Flexible Platform Archi-tecture for Developing Game NPC Intelligence with Load Sharing. In: The Workshop of 2005 International Conference on Computational Intelligence and Security, Xi'an, China (2005)
9. Gomez, F., Miikkulainen, R.: Incremental evolution of complex general behavior. Adaptive Behavior 5, 317–342 (1997)
10. Moriarty, D.E.: Symbiotic Evolution of Neural Networks in Sequential Decision Tasks. Technical Report AI, Austin, pp.97–257 (1997)

11. Yong, C.H., Miikkulainen, R.: Cooperative Coevolution of Multi-Agent Systems. Technical Report AI, pp. 01–287. University of Texas at Austin (2001)
12. SADE, http://klobouk.fsv.cvut.cz/~ondra/sade/sade.html
13. Holland, J.H.: Adaptation in natural and artificial systems. Internal Report. University of Michigan, Ann Arbor (1975)
14. García-Pedrajas, N., Ortiz-Boyer, D., Hervás-Martínez, C.: An alternative approach for neural network evolution with a genetic algorithm: Crossover by combinatorial optimization. Neural Networks 19(4), 514–528 (2006)
15. Stanley, K.O., Bryant, B.D., Miikkulainen, R.: Real-time neuroevolution in the NERO video game. IEEE Transactions on Evolutionary Computation 9(6), 653–668 (2005)
16. Buckland, M.: AI Techniques for game programming (2004)

Knowledge Discovery by Network Visualization

Hong Zhou[1], Yingcai Wu[1], Ming-Yuen Chan[1], Huamin Qu[1],
Zhengmao Xie[2], and Xiaoming Li[2]

[1] The Hong Kong University of Science and Technology
[2] Peking University
{wuyc,zhouhong,pazuchan,huamin}@cs.ust.hk,
{xzm,lxm}@pku.edu.cn

Abstract. Hyperlinks among webpages are very important information and are widely used for webpage clustering and webpage ranking. With the explosive growth in the number of webpages available online, the exploration of hyperlinks among webpages becomes a very challenging problem. Information visualization provides an effective way of visualizing hyperlinks and can help users gain insights into the relationships of webpages.

In this paper, we present some novel computer graphics techniques to visualize the hyperlinks among webpages. We propose a visual encoding scheme for five dimensional hyperlinks data and two constrained 3D layout techniques for incoming and outgoing links of a single webpage. To reveal the hierarchical structure of webpages as well as the hyperlinks information, we extend the treemap representation. Our representations are visually appealing and can effectively reveal linkage patterns among webpages. Experimental results and a user study demonstrate the effectiveness of our system. Our system can facilitate E-learning and help students understand the complex structures and hidden patterns in network datasets.

1 Introduction

Hyperlinks among webpages are important and useful features of the Internet. They can reveal interesting information about a webpage. For example, hyperlinks can be used for clustering webpages because relevant webpages are usually connected via hyperlinks. In addition, hyperlinks can be used to rank webpages. Many hyperlinks pointing to a webpage usually mean that this webpage is important. Recently, search engines have been relying more and more on linkage data to determine webpage rankings. Hyperlinks are also widely used in data mining and other applications. Sometimes, users want to manually explore hyperlinks data to reveal hidden linkage patterns associated with a webpage or a network. Simply reading the results from database queries may be tedious and inefficient. A picture is worth a thousand words. Computer graphics and imaging techniques have thus been introduced to help users explore linkage data. With the rapid growth in the number of webpages, visually exploring links among webpages has become a critically important technique. The linkage patterns also pose special challenges for students. It is difficult to understand the various attributes of webpages and their complicated relationships.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 240–251, 2008.

Webpages can be naturally organized into a hierarchical structure. Webpages are at the bottom level of this hierarchy. One level up is the webhost which may host many webpages. One institution (e.g., organization, company, university) can have more than one webhost. These institutions may belong to different domains. The visualization tasks for hyperlinks among webpages can be classified into three categories: visualizing links coming in or going out a single webpage; visualizing hierarchical structures (i.e., webpage → webhost → institution → domain) of webpages; and visualizing links among a group of webpages, webhosts, or institutions. Hyperlinks are usually visualized by using node-link diagrams, where nodes represent webpages and edges represent links. Effectiveness of the graphical representations and scalability of these representations for very large data are two major challenges for node-link diagrams. In this paper, we try to address these two issues for linkage data visualization.

We propose several novel visualization techniques to visualize hyperlinks among webpages. Specifically, we propose a visual encoding scheme which can visualize four dimensional data associated with incoming or outgoing hyperlinks of a single webpage or webhost. We use constrained 3D layouts to make our representation scalable for large data. We develop an enhanced treemap representation to visualize the hierarchical structure of a webpage or webhost along with associated hyperlinks information.

Our method is not limited to the visualization of hyperlinks among webpages. It can be further extended for more general network visualization problems. Hyperlinks among webpages are just an application of network visualization, which uses interactive computer graphics and imaging techniques to help users gain insights into massive data whose internal relationships can be described using networks or graphs. Some other examples include citations in scientific papers, airline routes, and social networks. The rapid growth in the size and complexity of these data have made network visualization a very important and challenging problem for information processing. Our methods can be applied to other network visualization problems. Experiments on a real dataset demonstrate that our system can facilitate knowledge discovery and help students find patterns in network datasets.

2 Related Work

Website Maps. There have been some works on website maps which visualize the linking relations among webpages in that website. This can help users navigate and search complex websites with the big picture of the website structure in mind and avoid getting lost during browsing. WebTracer[1], which is being developed by Tom Betts, is a tool for mapping the structure of websites. This freeware tool utilizes a 3D molecular model visualization to show hyperlinks of a given site. Spheres are used to represent webpages - the larger the sphere, the heavier the linkage to and from this page. The red and blue edges represent links between pages. Tree Studio, developed by Inxight Software (a spin off from Xerox PARC), provides a neat fisheye style interactive website map. There are various notable works on visualizing the structures and the evolution of websites and the Internet [1,2,3]. In this paper, we only focus on one important aspect of the Internet, i.e., the hyperlinks among webpages.

[1] http://www.nullpointer.co.uk/-/webtracer.htm

Links among Webpages and Network Visualization. The links among webpages are usually represented as either a matrix or a node-link diagram. The node-link diagram representation is more popular. Munzner [4] gave an excellent survey for network visualization. One of the earliest network visualization systems is SeeNet, which has had profound influences on subsequent works. The simple node-and-link map is used, in which the color and thickness of the lines represent the strength of the relationship while the glyphs encode statistics data associated with the nodes. Cox *et al.* further extended SeeNet to SeeNet3D [5] by exploiting more available space in 3D to decrease line intersections and reduce visual clutter. Such simple graphical primitives were amenable for slow computers with limited graphics capacity at that time. However, they have limited ability to encode multivariate information associated with network data and have poor scalability for massive data. Visual metaphors and interaction are two main factors affecting visual scalability [6] and many studies have been conducted in these two directions. Rafier and Curial [7] proposed an effective visualization method for large networks through sampling.

Treemaps. Treemaps are very effective graphical representations for data with hierarchical structures. In treemaps, the size and color of the individual rectangles are significant and can be used to encode data attributes. For example, if the tree represents a file system hierarchy, the size may be proportional to the size of the respective file and the color indicates the file type. The layout algorithms for the treemap and its various extensions have been thoroughly studied. An excellent survey can be found [8]. Some notable, variations of treemaps include Cushing treemaps, 3D treemaps [9,10], and Voronoi treemaps. An interesting technique has been introduced by Fekete [11] to visualize a graph as a treemap with overlaid links.

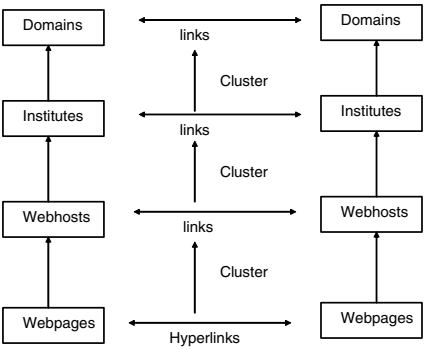


Fig. 1. Hierarchical structure of webpages

3 Data Collection and Preparation

The webpages used in this experiment were collected and processed in the network laboratory of our department. The webpages were first retrieved by crawlers and then hyperlinks information was extracted. After analyzing the webpages and with the help

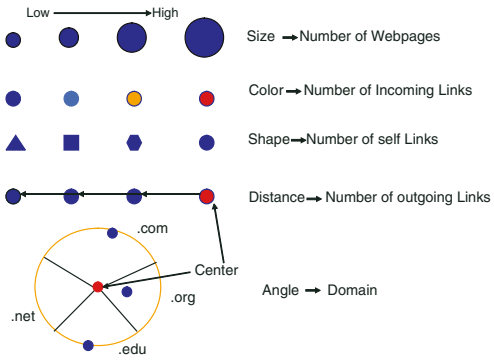


Fig. 2. Our visual encoding scheme

from some domain registration institutions, we constructed a four level hierarchical structure from the webpages, i.e., webhost (a server hosting webpages), institution (a physical institution which has one or more webhosts), and domain (e.g., .edu, .net). Figure 1 shows the hierarchical structure constructed from webpages.

To simplify the presentation, we introduce the following terms: *a node* represents either a webpage, a webhost, an institution (e.g., a company, an university, an organization), or a domain (e.g., .com, .net. .org); *the size of a node* represents the size of a webpage, the number of webpages in a webhost, the number of webhosts in an institution, or the number of institutions in a domain; *outgoing links of a node* represent all the links pointing to other nodes from this node; *incoming links of a node* represents all the links pointing to this node from other nodes; two nodes are called *connected* if there are hyerlinks between them; *the strength of a connection or a link* represents the number of links between two nodes.

For each node, we collected the following information: Node ID and name; Parent and children nodes of this node; Links among nodes; Size of the node; Number of incoming links; Number of outgoing links.

All these data were processed, indexed, and stored in a MySQL database for efficient query. Without loss of generality, all the nodes used in this paper are webhosts unless otherwise stated.

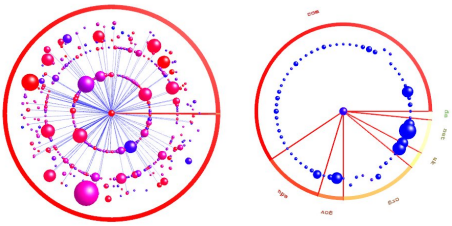
4 Incoming/Outgoing Links for One Node

To visualize linkage information for one single node, we can draw a simple graph or node-link diagram where a node represents a webpage, a webhost, an institution, or a domain and a link between two nodes indicates that they are connected via links. To distinguish the node (i.e., a sphere or other graphics primitive) used in the node-link diagram and the node (e.g., a web host) in the linkage data, we use the graph node and the web node, respectively, if ambiguity arises. In this section, we first introduce a graphical representation which can encode four dimensional linkage data for a single node. After

that, we present two techniques that can dramatically improve the effectiveness of our encoding scheme.

4.1 Encoding Scheme for 4D Data

For each node, we want to find out which nodes are connected to this node via incoming and outgoing links as well as some extra dimensional data attributes associated with these nodes and links. For links, we want to find out the type of links (i.e., incoming links or outgoing links) and the strength of links (i.e., number of links). For nodes, we want to know the size of the nodes, the total number of incoming links, and the domain of the nodes if the nodes are not domain (Optional).



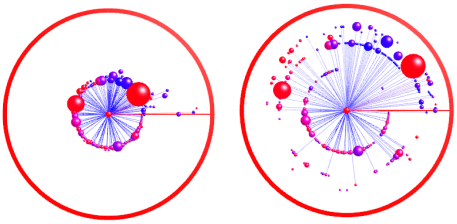
(a) (b)

Fig. 3. Visualizing incoming and outgoing links for one webhost

Encoding all these information into one image can help users find the possible correlations between different attributes of nodes and links. We design the following visual encoding scheme for five dimensional linkage data:

- The size of a web node is encoded using the size of the corresponding graph node.
- The number of outgoing links from the central web node to another web node is encoded using the distance between them. The larger the number of links, the closer the two nodes. The distance can naturally represent the strength of the relationship between two nodes.
- The number of all incoming links of a web node is encoded using the color of the corresponding graph node.
- The incoming links to the central node from an outside node are represented by the line between the two node. If there is no outgoing link from a web node to the central web node, no line will be drawn between them. The color of the lines/links encodes the number of outgoing links from the web node to the central node.
- Nodes are clustered according to their domains. For example, if two webpages belong to the same domain, they will cluster together. We use a pie-style graph (See Figure 3b) to show the percentage of each domain among all domains.

Figure 2 shows our visual encoding scheme. Figure 3 shows all the nodes connected to one single node via links and the associated statistics information using our 4D visual encoding scheme.



(a) (b)

Fig. 4. Distance histogram equalization: (a) before equalization; (b) after equalization

4.2 Distance Histogram Equalization

For very large datasets, our encoding scheme may not have enough space to layout so many nodes. Many nodes may overlap, which make exploration difficult. We can pre-cluster some nodes so that the number of nodes can be reduced to a manageable level. We introduce two techniques, *distance histogram equalization* and *constrained 3D layouts* to improve the layout of our methods and to display more nodes.

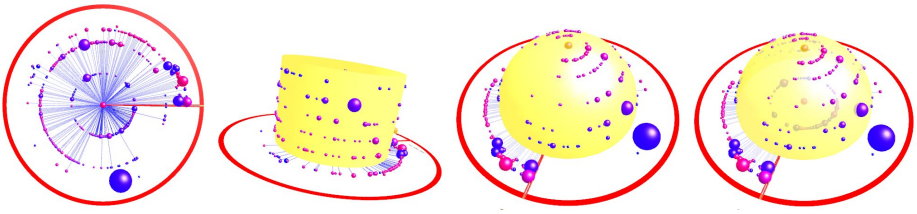
Our encoding scheme puts the nodes with similar connection strength in a similar distance from the central node. If the connection strength has a very uneven distribution, then many nodes will cluster on certain circles while other regions only have few nodes (See Figure 4a). To solve this problem, we introduce the **Distance histogram equalization**, which only maintains the relative connection strength for nodes. The idea is similar to image histogram equalization. Our method consists of three steps:

1. Distance quantization. The distance equals to the number of links. We first quantize the distance into a certain number of levels. The quantization is not linear. Because the nodes closer to the center node are usually more important, we give them more levels. If the nodes are far away, we just quantize a larger distance range into one level.
2. Distance histogram construction. For each distance level, we count the number of nodes falling into this distance range and then build a histogram of distance distributions for all nodes.
3. Distance histogram equalization. We apply histogram equalization to the distance histogram and then compute the new distance, and thus new position, for each node. Then all nodes will be displayed using the new positions.

Figure 4 shows the layouts before and after distance histogram equalization. From the figure, we can clearly see that the usage of space is dramatically improved.

4.3 Constrained 3D Layouts

To further make more nodes visible for users to explore, we propose two **Constrained 3D layouts**, the cylindrical view and the semi-spherical or dome view. We use a dome (i.e., semi-sphere) or a cylinder to show the nodes falling into the similar distance range. During an interaction process, users click on a node, then a cylinder or 3D semi-sphere will grow up and all the nodes falling into the same distance level with this node will



(a) (b) (c) (d)

Fig. 5. Constrained 3D layouts: (a) Original layout; (b) Cylindrical layout; (c) Semi-spherical or dome layout; (d) Transparent dome layout

be re-positioned on the surface of the semi-sphere or cylinder. Figure 5 shows the two 3D constrained layouts. Users can switch between cylindrical view and semi-spherical view. The transparency of the cylindrical surfaces and semi-spherical surfaces can also be adjusted. These layouts have the following advantages:

1. There is more space to layout the nodes. The nodes can be positioned onto a 2D surface instead of a 1D circle.
2. It is natural for users to understand that the nodes on the semi-spherical or cylindrical surface have a similar distance to the central node.
3. The constrained 3D layout can overcome some of the disadvantages of other general 3D layouts. One major disadvantage of using 3D layouts is the visibility and occlusion problem which may be confusing for some users. However, our constrained 3D layouts have the advantage that more space can be used but no serious occlusion problems will be caused.

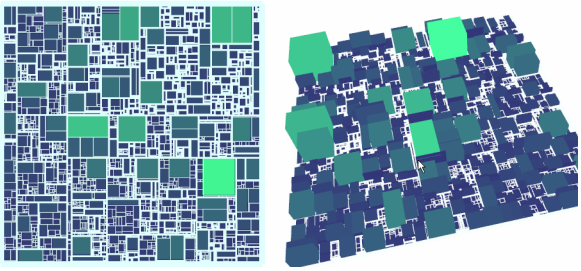
5 3D Treemap for Linkage Data

As mentioned earlier, webpages can be naturally organized into a hierarchy. This hierarchical structure may be useful for many applications. To display this hierarchical structure along with linkage information, we exploit the treemap representation. The treemap is a classic visualization technique to show data with hierarchical structures. Usually, treemaps can use the following features to encode three dimensional data into one image: the hierarchical structure, the size, and the color of the boxes.

In our example, we have more than three dimensional data. To encode the hyperlinks information with the associated multi-variate statistics data, we introduce a 3D treemap representation. Our method first shows a 2D traditional treemap. The hierarchical structure of boxes encodes the hierarchical structure of webpages. Users can then choose to use the size of the boxes and the color of the boxes to encode any two of the following four node attributes: size of this node, number of self links, number of outgoing links, or number of incoming links. Users can zoom in/zoom out the treemap and click on any box then the upper level boxes will be highlighted and their associated node information will be shown on a separate message window. In addition, after a user clicks on a box which represents a node, all other nodes connected with this node will be drawn using

3D boxes, where the height of the boxes can encode another statistic data attribute (see Figure 6b).

Compared with traditional treemap representations, our 3D treemap can encode two more node attributes, i.e., the relationship of nodes represented as the dimension of boxes (3D vs. 2D) and another attribute represented as the height of the box. The 3D treemap has been used before to visualize file system hierarchies by Bladh et al. [9]. They used the height of the box to encode the depth in the file tree. To the best of our knowledge, it is the first time that the 3D treemap is used to encode linkage data. Figure 6 compares the traditional 2D treemap and our 3D treemap.



(a) (b)

Fig. 6. Treemap for linkage data: (a) 2D treemap encoding 3D data; (b) 3D treemap encoding 4D data

6 User Study

A real system has been developed based on our methods to visualize the linkage data we collected. The system is installed on a Pentium(R) 3.2GHz PC with 1GB RAM and a 256MB Nvidia Geforce 6800 Ultra graphics card. MySQL is used to handle queries from users. To test the effectiveness of our encoding schemes and the usability of our system, we conducted a user study.

6.1 Procedures

The data we used consists of all the webpages in the "*edu.cn*" domain collected at Peking University. We obtained about 820 Chinese educational institutions and 4,333 webhosts. Ten students with normal color vision, which consist of five undergraduate students and five post-graduate students in the computer science department of Hong Kong University of Science and Technology, participated in our user study. All are experienced computer users but have no prior experience with our system or any similar information visualization systems. Before starting the user study, each user received a fifteen-minute training about the visualization and interaction techniques of our system. They were also provided a diagram similar to Figure 2, which shows our visual encoding scheme. After the training, each participant was asked to perform four tasks. The first task was designed to test the usability of our system. The participants were asked to do conjunctive queries using our system. The other three tasks were designed to test the

effectiveness of our visual encoding schemes. Ten subjects are divided into two groups. These two groups then explored the data using different encoding schemes (i.e., layouts with and without distance histogram equalization; 2D layout *v.s.* 3D constrained layout; straight-line layout *v.s.* curve layout). The four tasks procedures are listed as follows.

Task 1. We chose one specific webhost (*www.pku.edu.cn*) and then asked each subject to find out one or more nodes which are connected to this webhost and have : 1) a large number of webpages; 2) a large number of incoming links; and 3) relatively weak connection to *www.pku.edu.cn*. It is used to test the encoding scheme in Section 4.1. After finishing this task, each subject was asked about their feelings.

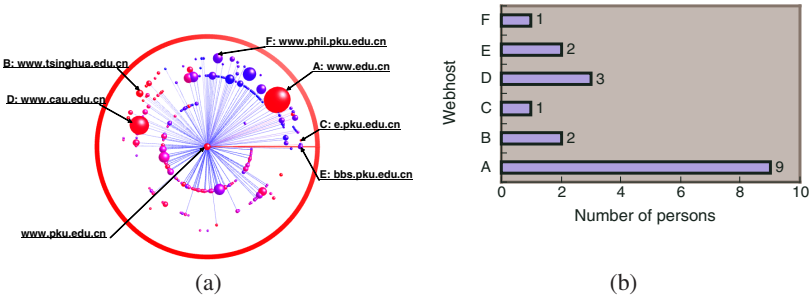
Task 2 & 3. We randomly divided the 10 subjects into two groups of five, to test our distance histogram equalization and constrained 3D layouts techniques.

In *Task 2*, the systems with and without distance histogram equalization were used by group A and group B respectively. Both groups were asked to perform a task similar to the first task, but the chosen webhost *www.pku.edu.cn* was changed to webhost *www.edu.cn*, and the three conditions were also changed slightly.

In *Task 3*, group A was asked to use only the 2D layouts, while group B was asked to use the 3D layouts. Distance histogram equalization was used by both groups. The question was still the same, but the chosen webhost was changed to *www.tsinghua.edu.cn* and the three conditions were changed slightly again.

After finishing task 2 and 3, each participant was asked about their experiences of answering these two questions.

After finishing all the three tasks, each subject was asked to play with our system for at least ten minutes. We observed their usage of our system and interviewed them about their experience and opinions after their exploration.



(a) (b)

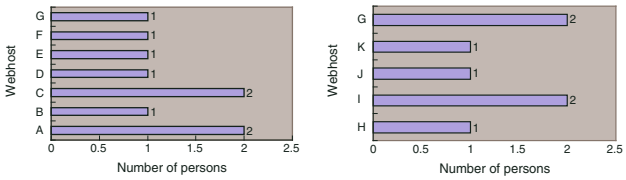
Fig. 7. Results for Task 1: (a) The linkage pattern for *www.pku.edu.cn* and the six webhosts chosen by the users; (b) The distribution of the user answers

6.2 Results and Discussions

The results of the user study tasks are summarized as follows.

Task 1. The average response time is 12.3 seconds. Ten subjects found out totally six webhosts, denoted by A-F (See Figure 7). From the bar graph (See Figure 7(b)), we can see that the answers overlap in some sense. Five of them pointed out that this question is very subjective, and another three persons considered it easy to answer.

Basically our system is good at those subjective tasks because users can easily get the overview of all the attribute values simultaneously. To formulate these three conditions using conjunctive queries in SQL is difficult because the large, hot and strong connections are relative and highly depend on the context. It is hard to decide the thresholds for the SQL queries. The results are also subjective and need human judgment. But this information is important for knowledge discovery and may reveal some patterns about the web data. A good solution is to show the overall distributions of all the attributes together and then let users to choose the answer using their own judgment. Our multivariate data encoding scheme can facilitate this process and help users for knowledge discovery.



(a) (b)

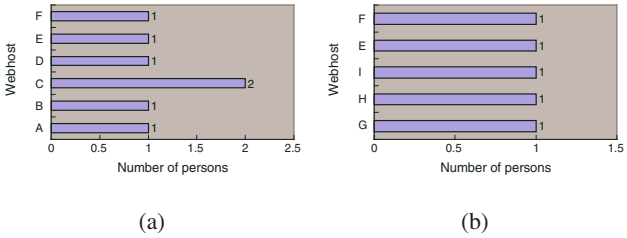
Fig. 8. Results for Task 2: (a) Result from group A. The average response time is 39.6 seconds; (b) Result from group B. The average response time is 21 seconds.

Task 2 & 3. For *Task 2*, the results are showed in Figure 8. The average response time of group A is about two times longer than that of group B. Only webhost *G* (See Figure 8) was pointed out by both groups, and more webhosts are pointed out by group A than by group B. From Figure 4, we can see that with the distance histogram equalization the usage of space is dramatically improved. Thus, the users can easily shrink the range of candidates and find out the final results faster after the equalization. In addition, the answers may be more reasonable and accurate.

For *Task 3*, the results are showed in Figure 9. The average response time of group B is about three times longer than that of group A, because the subjects in group B spent more time on 3D rotation and exploration. After exploring the constrained 3D layouts, each subject of group B chose only one webhost while two group A members picked out multiple choices.

For the question about their experiences of answering these two questions, two group A members disliked the layout (i.e., without distance histogram equalization) in task 2. They complained that too many nodes were clustered together. One group B member said that the cylindrical layout was very cool and really helpful. However, another person thought that using the 2D layouts was easier than using the 3D layouts to find the answer because of the longer time spent on the rotation.

Most participants mainly used the visualization scheme presented in Section 4.1 to explore the data. They are fascinated by our system. When playing with our system, most users can find some interesting information and some emitted comments like “Oh, my favorite college is so blue!” and “Why don’t these two universities have many links between them?” They were usually attracted to some noticeable but unknown webhosts. Some of them even opened the web browser to take a look at these webhosts.



(a) (b)

Fig. 9. Results for Task 3: (a) Result from group A. The average response time is 36.2 seconds; (b) Result from group B. The average response time is 86 seconds.

One example is a Chinese educational portal webhost (i.e., *www.edu.cn*), which was noticed by all users, because this webhost is very big, red, and nearly all the Chinese educational institutions are linked to it. Another webhost discovered by most subjects is *www.gaokao.edu.cn*, which is a very large and blue webhost. Most subjects are not aware of the existence of these webhosts before the user study. Some unexpected linkage patterns were also found out by our users. For example, someone commented that the big and red nodes are usually far away from the central web node. For example, in Figure 5a, 7a, and 3a, many big and red nodes are not very close to the central web node. We found that this is also a common fact for many other webhosts. This is an interesting finding, which means, among the Chinese education institutions, many webhosts are connected to some big and hot webhosts but their connection strengths are not as high as we expected. We were surprised by their findings because as developers we did not notice these patterns before.

The 3D treemap display was also played by some participants. They liked to change the encoding scheme of treemap frequently and click those attractive boxes to see what they are. One user pointed out that nearly all the institutions contain a BBS webhost (e.g., The University of Science and Technology of China has *bbs.ustc.edu.cn*) which has a large number of webpages and many incoming links.

7 Conclusion

In this paper, we introduced some novel computer graphics techniques for visualizing hyperlinks among webpages. We proposed a 4D encoding scheme and constrained 3D layouts to visualize incoming/outgoing links for a single webhost. We further extended the treemap representation so the links information can be shown along with the hierarchical structure of webpages. A user study was conducted to demonstrate the effectiveness of our encoding schemes and the usability of our system. Some unexpected patterns have been detected by students using our system.

Our current work only focuses on effective visualization techniques for medium size data. With the improvement of data collection techniques, more data such as the click stream for hyperlinks and history of webpages will be available for users to explore. How to visualize large data and encode more than five dimensional data is worth further study. We also plan to extend our methods for general network visualization.

Acknowledgments

This work was supported by Hong Kong RGC grant CERG 618706 and NSFC grant 60573166.

References

1. Cheswick, B., Burch, H., Branigan, S.: Mapping and visualizing the internet. USENIX, pp. 1–12 (2000)
2. Chi, E.H., Pitkow, J., Mackinlay, J., Pirolli, P., Gossweiler, R., Card, S.K.: Visualizing the evolution of web ecologies. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 400–407 (1998)
3. Chi, E.H., Card, S.K.: Sensemaking of evolving web sites using visualization spreadsheets. In: Proceedings of the 1999 IEEE Symposium on Information Visualization, pp. 18–25 (1999)
4. Munzner, T.: Interactive visualization of large graphs and networks. Ph.D. dissertation, Stanford University (2000)
5. Cox, K., Eick, S., He, T.: 3d geographics network displays. SIGMOD Record 25(4), 50–54 (1996)
6. Eick, S., Karr, A.: Visual scalability. Journal of Computational and Graphical Statistics 11(1), 22–43 (2002)
7. Rafiei, D., Curial, S.: Effectively visualizing large networks through sampling. IEEE Visualization, 48–55 (2005)
8. Shneiderman, B.: Treemaps for space-constrained visualization of hierarchies, http://www.cs.umd.edu/hcil/treemap-history/
9. Bladh, T., Carr, D., Scholl, J.: Extending tree-maps to three dimensions: A comparative study. In: Proceedings of 6th Asia Pacific Conference on Computer Human Interaction, pp. 50–59 (2004)
10. Tanaka, Y., Okada, Y., Niijima, K.: Treecube: Visualization tool for browsing 3d multimedia data. In: Proceedings of Information Visualization 2003, pp. 427–432 (2003)
11. Fekete, J.D.: The infovis toolkit. In: IEEE Symposium on Information Visualization, pp. 167–174 (2004)

Research on Emotional Vocabulary-Driven Personalized Music Retrieval

Bin Zhu[1] and Tao Liu[2]

[1] College of information science and technology, Zhejiang Shuren University, 310015, Hangzhou, Zhejiang Province, P.R. China
zhubinxx@sohu.com
[2] College of computer science and technology, Zhejiang University, 310027, Hangzhou, Zhejiang Province, P.R. China

Abstract. In this paper, aimed at emotional interactive needs in music retrieval, a new retrieval method, based on dynamic characteristics of fuzzy psychological and linguistic value computing of music emotion, is putting forward. And an improved interactive genetic algorithm is designed for the emotional retrieval goal. In order to reduce users' fatigue of retrieval process, choice operator for fine breed, crossover operator holding the dominant element of emotion vector, and self-genetic operation strategy are designed.

Keywords: Music retrieval, Linguistic value model, IGA, Genetic operation.

1 Introduction

With the rapid development of multimedia and in-depth popularity of Internet technologies, the music information is growing rapidly. How effective retrieval for music is becoming an important research area in modern information retrieval. A typical traditional practice is to use text formats such as names of songs, singers or lyrics keywords, as users must remember that relevant information. However music is wildly different with text data, the user's needs are often based on the music semantics as melodies, form, style and emotion, and so on. In the means of semantic-based retrieval, the user only needs to present the musical content of a semantic description, for example: some music samples, such as a user playing or singing the melody repeat [1], a school or style (classical, rock, etc.), or some kind of emotional acoustic sensing features (sorrow, joy, etc.).

Different retrieval conditions of semantic information are corresponding to different music retrieval technology. After the pioneering papers [1] in ACM multimedia seminar, a considerable amount of research on "Query by Humming" retrieval system is reported, which adopts a more natural way of interaction, so that users only need to remember songs melody.

In the study of music retrieval based on the style or emotion, the premise work is organizing the music in semantic information. Island of Music System [2] introduces a method based on the similarity of perception and reduced-dimension to extract rhythm patterns, use self-organizing map (SOM) data analysis methods for data

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 252–261, 2008.

classification, and then use smoothing data histogram (SDH) for visualization and browsing. This system is particularly suited to the music retrieval requests such as "sounds like how" and "what emotional color with the works". However it is a pity that the system itself does not provide the retrieval application.

In article [3] a pioneering work in emotion-driven music retrieval is presented, which implements the retrieval mode in 200 pieces of MIDI music database based on the binary encoding interactive genetic algorithm, whose genetic operating is evolved in the music characteristic space, the user's choice of emotional needs is integrated into the interactive process. However, in the retrieval process, users' fatigues are occurred easily, and the emotional expression model is also not match with the psychology research.

Juslin's researches [4] on music and its emotion connotation is much representative, but there is little discussion about music retrieval based on emotional information. This paper is studying on the emotional retrieval method of music, which adopts emotional vocabulary as input conditions.

2 Linguistic Value Model for Music Emotion

Since music emotion behaves as a special linguistic value system, we get the following definitions [5].

Definition 1 (Linguistic value model): Binary variable $< LA, R >$ means that the linguistic value model:

$$LA = \{L_1, L_2, ..., L_n\} \tag{1}$$

$$R = (r_{ij})_{n \times n}, r_{ij} \in [0,1], i, j = 1, 2, ..., n \tag{2}$$

LA is a class of limited linguistic values, and R is defined as fuzzy similarity on LA, and n is the number of elements in classes. rij is a repetition degree of semantic between Li and Lj. Clearly, fuzzy relation matrix R is a symmetry matrix, meet with two conditions: rij= rji and rii=1.As figure 1 shows, the eight subclass of emotion space is expressed with eight elementary adjective as a ring named as Hevner emotion ring: LAoM={ Dignified, Sad, Dreamy, Soothing, Graceful, Joyous, Exciting, Vigorous}, it is noted: LAoM ={LAoMi,i =1,...,8} .

According to the definition of linguistic value of music emotion, by expanding the concept of semantic similarity, we can easily get music emotional vector definition as follows:

Definition 2 (music emotion vector): As to a certain music, noted as Mc, with independent semantic emotion, its emotional connotation is defined as eight-dimensional vector, signed as E, in Hevner Emotion Ring, each of whose elements, noted as e_i , performs for the semantic similarity relations of music and each emotional linguistic value, with 0-1 expression. And we called the Vector as music emotion vector, signed as E:

$$E = (r(M, LAoM_1), ..., r(M, LAoM_i), ..., r(M, LAoM_8)) \tag{3}$$

The sub-emotion corresponding to the max numerical is called as the dominant emotion for music.

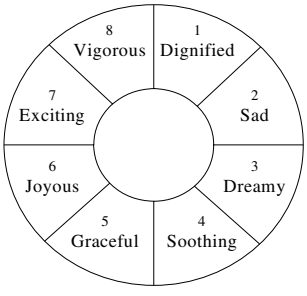


Fig. 1. Hevner Emotion Ring

For instance, after some reasoning process, the emotional connotation of a certain music M can be expressed as follows: (0.2, 0.6, 0.9, 0.4, 0.3, 0.1, 0.0, 0.1), and so its dominant emotion is "longing", with emotional semantic similarity value is 0.9, while the music also includes a "sad" emotional connotation, with a little lower value as 0.6, which can be termed as the secondary dominant emotion. Anyway, the emotional connotation of music M can be described as "very longing, and some sadness".

The preliminary work of music retrieval is automatic identification and emotional machinery tagging for music clips database. Figure 2 shows the emotional music of tagging and database building process [6]. First, we build the music clips database with independent emotional semantic through music feature recognition, and using the theme melody to represent the music as abstract. And then we tag music clips with emotion vector based on the IGEP algorithm, which achieves mapping from a 10-dimensional feature space to the 8-dimensional emotional space.

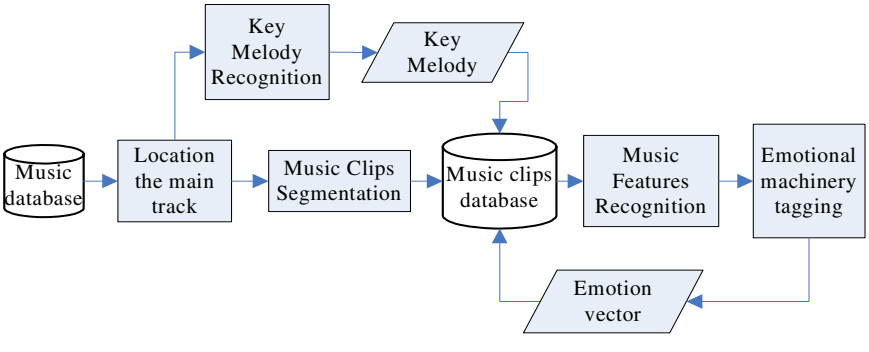


Fig. 2. Machinery tagging process for music emotion

3 Personalized Emotional Music Retrieval

In the means of emotional music retrieval, users' needs may be difficult to determine with the decided vocabulary to express or the needs is very vague, which needs the

help of outside advice. In all, the needs of user are uncertain, as we call this mode of retrieval as the personalized retrieval. As the communication among peoples, only interactive process helps to mutual understanding and tacit agreement. As shown in figure 3, interactive genetic algorithm (IGA) require user to evaluate the satisfactory degree of the newly- evolutional individual through the interactive process, which is used to measure its distance with psychological space, and then the system performs subsequent operation and produce a new generation of groups by the corresponding fitness function, and constantly repeat to get an optimize individual meeting the user's requirements. The purpose of interactive process is to integrate the human wisdom and computer technology for improving the performance of genetic algorithms. Since the preference, intuitive, emotional and other psychological factors of users can be identified as a fitness value of the characteristics that can be binding to the target system, the IGA algorithm is also applied in the music retrieval [3], [7].

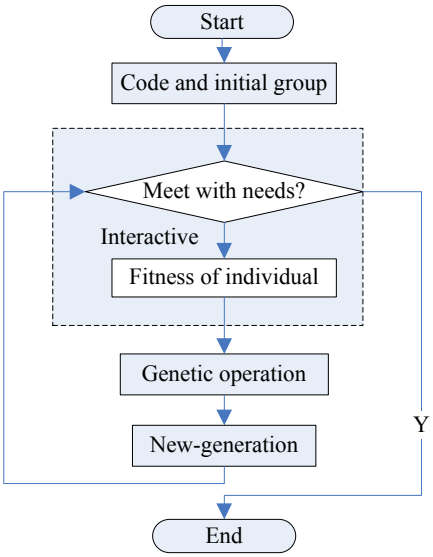


Fig. 3. Flowchart of IGA

The same as the genetic algorithm, IGA also requires fewer individuals of each group and more minimized evolutionary generation as possible, in order to speed up the retrieval speed because of music's timing characteristic, with which a person's memory is limited therefore the presented individual must be restricted. According to the same reasoning, evolution generation must also be limited.

The most serious problem of IGA is how to reduce fatigue in the interactive progress. Since people should communicate with computer, giving assessed value of each individual in each generation, larger the number of individuals, longer evolutionary time, more easily result in fatigue, especially in the case of facing these individuals with little difference and difficulty to distinguish, which will induce greater psychological pressure and thus more easily produce fatigue.

3.1 Coding Process and the Initial Generation

In IGA for emotional music retrieval, not only the specific characteristics of emotional linguistic value of music but also some specific genetic operators should be taken into account. Since music clips database has been established through the emotional tagging process, the supervenient music retrieval can be performed directly in the music emotional space.

By now, the commonly used method of coding is the binary-coded and Real-coded. According to the definition of emotion vector, this thesis is based on the real number of eight dimensional vector, ranging between [0, 1], so the higher accuracy is required (5 location after decimal point). If the binary code is adapted, its length of solution is 8 * 17 = 136, the facing search space is 2^{136} , and that is very big. In addition, the use of Real-coded approach can also help development of the specialized genetic operator for music retrieval. Therefore, this paper uses Real-coded method [8].

Similar with the definition of emotion vector, style of gene series $X_i(i = 1, 2, ...N)$ is expressed as $(a_{i1}, a_{i2}, ..., a_{in})$, in which we have $a_k \in D(a_k) \subset R(k = 1, 2, ..., n)$, n = 8 for the dimension of space, and N for the population size (this paper evaluated as 8).

For the average retrieval process, the selection of first generation is very important. The retrieval process will be speed up when the music has certain characteristics consistent with the user's needs. For the greatest possible diversity of the sample, the system chose one from eight pieces of music whose value of similarity with dominant emotion is biggest as the first generation sample groups. Those eight pieces of music would certainly include the characteristics of user's demand, so it is bound to increase the efficiency of retrieval, but also diversity of samples can be insured.

3.2 Genetic Operator

Genetic operations play vital role in IGA, which achieve the survival of the fittest of the evolutionary process by genetic operators (such as choice operator, crossover operator and mutation operator, etc.) and the degree of adaptation to the environment.

1. Choice operator for fine breed

To ensure individual with greater fitness will be able to be retained in the next generation, while accelerating the process of the convergence of GA, the evaluation of music achieved by users will be cognized as the criterion for choice to choose 8 subchromosomes. The calculation process is as follows:

(1). Consider numerical evaluation of the users as the coefficients corresponding to emotional vector of each individual, and calculate the reference emotional vector according to multi-source information fusion method based on music emotional linguistic value model [5];

(2). Calculate the similarity between the emotion vector of each music individual sample and reference emotional vector, and evaluate fitness value fi of this individual i, and arrange those individuals by descending of fitness values;

(3). Calculate the fitness ratio of individual i according to the formula $P_{si} = f_i \Big/ \sum_{i=1}^{N} f_i$;

(4). Calculate the individual number expected in next generation: $m_i = NP_{si}$;

(5). Identify existent number of each individual in next generation by the integral part $\lfloor m_i \rfloor$ of m_i;

(6). Arrange individual by its fitness ratio, and chose $N - \sum_{i=1}^{N} \lfloor m_i \rfloor$ individual into the next generation until the number of this group reached N.

2. Crossover operator holding the dominant emotion

In this paper, crossover operator holding the dominant emotion is adopted, by which the leading cognitive components of users can be maintained as far as possible.

Supposed $s_1 = (v_1^{(1)}, v_2^{(1)}, ..., v_n^{(1)})$ and $s_2 = (v_1^{(2)}, v_2^{(2)}, ..., v_n^{(2)})$ as the father of two vectors, $s_z = (z_1, z_2, ..., z_n)$ and $s_w = (w_1, w_2, ..., w_n)$ as two children obtained by crossover operation.

Firstly, algorithm ranges each element of the father vector; secondly, selects the k component with small quantity; lastly, generate $n-k$ random number range of $(0,1)$, noted as $\alpha_{k+1}, \alpha_{k+2}, ..., \alpha_n$. In order to illustrate the problem, without losing the general assumptions, supposed the last k component is the smaller, and then two child generations will be defined as:

$$s_z = (v_1^{(1)}, ..., v_k^{(1)}, \alpha_{k+1} v_k^{(1)} + (1-\alpha_{k+1}) v_k^{(2)}, ..., \alpha_n v_n^{(1)} + (1-\alpha_n) v_n^{(2)}) \quad (4)$$

$$s_w = (v_1^{(2)}, ..., v_k^{(2)}, \alpha_{k+1} v_k^{(2)} + (1-\alpha_{k+1}) v_k^{(1)}, ..., \alpha_n v_n^{(2)} + (1-\alpha_n) v_n^{(1)}) \quad (5)$$

Of course, here we can take $\alpha_{k+1} = ... = \alpha_n$, thus there only needs one random number, and the value k is considered as 4-6 generally.

3. Mutation operator

Mutation operator restores the diversity of group through changing some gene of each individual randomly, which makes the retrieval able to achieve the entire solution space. In real-coded case, the mutation operator does not play a role as binary code, which only restores the loss of diversity of group simply; otherwise it has become a major retrieval operator. This paper adopts normality mutation operator used in the genetic algorithm frequently.

4. Self-genetic operator

In IGA of this paper, a self-genetic operator on the foundation of music emotion linguistic value model is designed to take full advantage of interactive information to reduce mental fatigue, by which system can chose the next generation automatically

through parametric curve interpolation algorithm based on the historical records of reference vector of every ancient generation when users feel fatigue.

Supposed system records the four emotional reference vectors $X_i = (x_{ij}), i = 1,...,4, j = 1,...,8$, and j represents that each emotion component, the three curves in this 8-dimensional space can be expressed as:

$$x_j = a_j + b_j t + c_j t^2 + d_j t^3, j = 1,...,8$$

The parameter $t = 0,...,3,...$ notes that the recent evolution number.

In addition, similarity between every two reference vector can be calculated as $S_{ik}, i, k = 1,...,4$ and $i \neq k$, then the general similarity metric of all reference vector is $S_i = \sum_{k=1,k\neq i}^{4} S_{ik}$. By using these four similarities as counterpart contribution on the curve interpolation corresponding to every vertex, we get the matrix expression of the curve:

$$P = P(t) = [1 \quad t \quad t^2 \quad t^3] \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{11}{6} & 1 & -\frac{3}{2} & \frac{1}{3} \\ 1 & \frac{1}{2} & 2 & -\frac{1}{2} \\ -\frac{1}{6} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} S_1 X_1 \\ S_2 X_2 \\ S_3 X_3 \\ S_4 X_4 \end{bmatrix} \tag{6}$$

And then the interpolation curve can be calculated by using the four reference vector to predict the next generation (t = 4) emotional reference vector. After the fitness value of new generation are calculated by system instead of user's interaction according to the formula above, the autonomous GA process will be restart.

3.3 Method of Identifying New Generation and the Termination Conditions of Evolution

After a series of genetic operation, the chromosome of group has been changed for some new type of chromosome produced. Since limitations of storage capacity of music clips database results that there may exit no be new corresponding items of chromosome in the database. This paper produces a new generation by recently neighbors Law.

Since evolutionary termination conditions have sth to with the quality of the solution, if we adopt a pre-determined the number of generation as a termination conditions, the complexity and convergence of the algorithm are difficult to control because different users have different cognitive ability, so it will be very difficult set a reasonable number of generation. In this paper, when the majority of music clips which users feel satisfied (generally preferable for 5) exit, the system identify that users have found that the needs of music and terminate operations.

3.4 Experiment

This method does not require users to have a prior clear impression of goal music and the user only has to select and score the music represented. In the interactive process, users can change the target music demands at any time. The whole process is described as follows:

A) Produce the initial individual groups automatically;

B) Evaluate the individual sample according to the relevance of the goal subject: users choose some musicians (above two) to positive or negative assessment and the fitness values of each individual will be calculated;

C) Check if the termination conditions of evolution are meeting: N to D), Y to the end;

D) Genetic operate on the group according to the fitness values calculate by step B): implement choice, crossover and mutation operation, and produce a new generation of individuals. The crossover probability defaults to 0.3, which users can change at any time;

E) Check the user fatigue and the number of generation, and if users feel fatigue and the number of generation has more than four, system goes to step F), the self-genetic processes, or to B);

E) Calculate the fitness value of reference vector according to the records of outstanding individual evaluated by users;

F) Calculate the fitness values of each individual of the generation and go to step C).

In the algorithm, with the deepening of the process of evolution, the target music chosen by the user should be increasingly clear. And so we can have better judge rules, and get more satisfactory results for users.

In our experiment, every generation provided eight music clips to users, and users give the evaluation of satisfaction about every song based on his or her audition.

We invite six boys and six girls to our experiment and divided them into six groups to test the algorithm contained absolute of entire operators designed above in a database including 1,190 musicians, and the goal music emotional linguistic words are

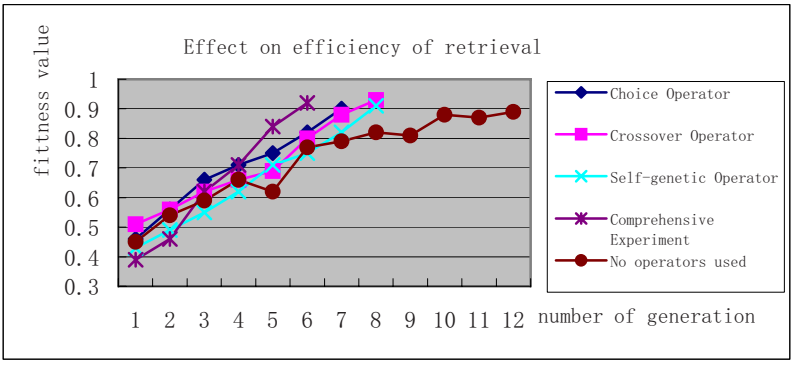


Fig. 4. Genetic operators

"romantic", "fantasy" and "sanctity", "dark", "anger" and "surprised". If the number of generation reaches 15, the retrieval is considered failure.

Figure 4 gives the average number of generation which users find the satisfactory music. These experiments show that the genetic operators designed above can effectively accelerate the convergence rate, so that we can better reduce the user's fatigue.

From the experimental results, the ratio of music clips in each generation meeting with the needs of user can be gradually increased through the interactive process. In all operators, the contribution of choice operator for the improvement of genetic convergence is the most notable. And under the operators synthetically, algorithm can effectively accelerate convergence, and find the music matching the emotional needs of the users.

4 Conclusions

In this paper, several major works are as follows:

(1) To identify and tag for emotional connotation of music clips based on linguistic value model of music emotion automatically;

(2) To design an improved interactive genetic algorithm against personalized retrieval demand, and emotional language-driven music retrieval;

(3) To design a series of strategies in order to effectively accelerate convergence and eliminate user's fatigue produced in the interactive progress, such as: choice operator for fine breed, crossover operator holding the dominant element of emotion vector, and self-genetic operation strategy.

Experiments show that, research in this thesis establishes an emotional bridge in the Human-Computer interaction of music retrieval. Computer understands the emotional needs of users in the in the interactive process while users can find the music to meet their own needs.

References

1. Ghias, A., et al.: Query by humming: Musical information retrieval in an audio database. In: Proc. of ACM Multimedia 1995, November 1995, pp. 231–236 (1995)
2. Pampalk, E.: Islands of music analysis, organization, and visualization of music archives. OGAI Journal (Oesterreichische Gesellschaft fuer Artificial Intelligence) 22(4), 20–23 (2003)
3. Cho, S.-B.: Emotional image and musical information retrieval with interactive genetic algorithm. Proceedings of the IEEE 92(4), 702–711 (2004)
4. Juslin, P.N., Friberg, A., Bresin, R.: Toward a computational model of expression in music performance: The GERM model. Musicae Scientiae, Special Issue 2001-2002, pp. 63–122 (2002)
5. Sun, S., Wang, X., Liu, T., Tang, Y.: Study on linguistic computing for music emotion (in Chinese). Journal of Beijing University of Posts and Telecommunications 29 (z2), 35–40 (2006)

6. Yang, C., Sun, S., Zhang, K., Liu, T.: Study on music emotion cognition model based on applying the improved gene expression programming. Journal of computational information systems (to be accepted)
7. Wang, S., Chen, E., Wang, S., Wang, X.: A Kansei Image Retrieval Based on Emotion Model (in Chinese). Journal of Circuits and Systems (6)(2003)
8. Herrera, F., Lozano, M., Verdegay, J.L.: Tackling Real-Coded Genetic Algorithms: Operators and tools for the Behaviour Analysis. Artificial Intelligence Review (12), 265–319 (1998)

Research on Update Service in Learning Resources Management System

Yongjun Jing[1], Shaochun Zhong[1,*], Jie Jian[1], and Xin Li[2]

[1] Institute of Ideal Information Technology,
Northeast Normal University, Changchun, 130024, P.R. China
sczhong@yahoo.com.cn
[2] Department of Computer Science and Technology,
Qingdao Agriculture University, Qingdao, 266109, P.R. China
jingyj873@nenu.edu.cn

Abstract. Resources update service is a pivotal problem in the construction of learning resources and the development of e-learning. Aiming at this problem, this paper presents a Distributed Push System of Learning Resources (DPSLR). The schedule between servers and users is a M ulti-objective Optimization Problem (MOP) in this system. We discuss the scheduling model of servers and users, and use an Ant Colony Optimization (ACO) to solve the MOP. We have developed a model to estimate the efficiency of the new resources update service.

Keywords: learning resources management, resources update; task schedule, MOP, ACO.

1 Introduction

The construction of learning resources is a systems engineering which includes design, development, usage, management and evaluation. It is not a one-step but a circular and dynamic process of *lack-supply-balance-lack-supply*. Resources update service is a necessary and significant service in learning recourses management system [1]. Since new resources are accumulating continuously and users are expanding rapidly, an obvious problem appears that users urgently ask for new resources while new resources are not timely available to the user. There are two resources update patterns at the present. One is the new resources disks delivered by people or mailed to users; the other is a modern update pattern that the new resources are published on Web, and users can download them. The former requires a long update cycle and wastes lots of manpower, material resources and finances. As the construction of informatization foundation has been consummating and Internet has been widely used, more researches focus on the latter [2].

In order to realize resources share effectively and resources update in time, we designed a Distributed Push System of Learning Resources, called DPSLR. In this system, we use ACO to resolve the schedule between servers and users. It has been

* Corresponding author.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 262–269, 2008.

proved that DPSLR was an effective Internet-based update service and solved disadvantages of traditional update pattern.

2 System Design

Fig. 1 shows the framework of DPSLR. In this system, there are several equal and distributed servers locating in different place. Every server can support update service for users independently under the control of the schedule server. Users and distributed servers are both managed by the schedule server. It is fit for global system control. In addition, the system has expansibility and fault-tolerant capability. When a new server is added, it is only need to register its information in the schedule server, and then the new server can work. When a server breaks down, the schedule server will automatically detect the error and collaborate with the users' requests again.

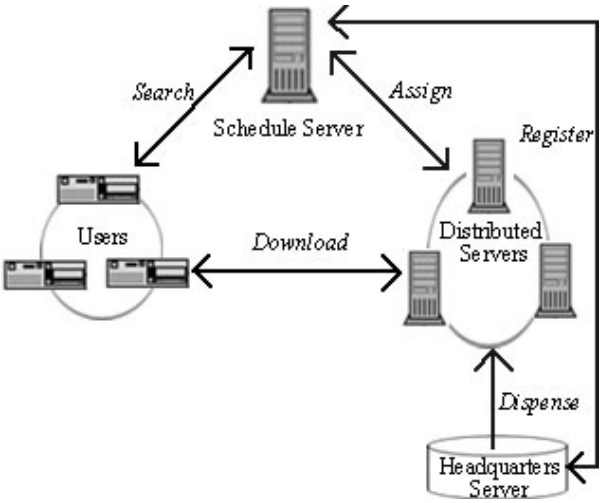


Fig. 1. Framework of DPSLR

Due to the dynamic characteristic of network communication and the requirements of the system performance, a system-oriented and dynamic schedule strategy is adopted in DPSLR [3]. We try to balance every server's load. It will be modulated during the running course of the system in order to reach the system load balance.

There are four parts (a headquarters server, distributed servers, users and a schedule server) and five operations (dispense, register, search, assign and download) in the system.

The working flows of the system are as follows:

(1) The headquarters server dispenses update-package data to distributed servers and registers description of update-package in the schedule server.
(2) The users search the update information from the schedule server.
(3) The schedule server assigns distributed servers to respond to the users' requests.
(4) The users download update-package data from distributed servers.

The functions of the four parts are as follows:

In the headquarters server, the update-package releasing system automatically compresses, encrypts and packs the new learning resources according to the schedule of resources development and the accumulation of resources. After getting distributed servers list from the schedule server, the headquarters server dispenses update-package data to the distributed servers and registers description of update-package in the schedule server.

In users' endpoints, the auto-update system is set up on their computers in the form of windows service. It searches update information from the schedule server according to the rules configured by users and downloads the data from the distributed servers returned by the schedule server. After the data is downloaded, the auto-update system will decompress, decrypt and unpack automatically. So far, the local resources will be updated and the upgrade will complete.

Distributed servers store update-package data and publish them by the protocols of HTTP or FTP. They supply download service for users under the control of the schedule sever.

The schedule server supplies the access interface for users, distributed servers and headquarters server in the form of web service, and stores information about the description of update-package, users and distributed servers. It also manages the match between users and distributed servers and assigns users' requests to related distributed servers. The schedule server is the core part of DPSLR. We will discuss it in details.

3 Schedule Server

3.1 Schedule Server Design

The Schedule Server is composed of task receiver, task scheduler, task dispatcher, task inspector, original task queen, optimized task queen, task log and database. Its framework is as Fig. 2.

The working flows of schedule server are as follows:

(1) The task receiver receives users' requests and collects related information to form the original task queen.
(2) The task scheduler re-arranged to form the optimized task queen using some optimization algorithm according to the original task queen and database.
(3) The task dispatcher assigns the related distributed servers to respond to users' requests according to the optimized task queen.
(4) The task inspector adjusts schedule strategy and re-optimizes task queen according to task log.

From a macroscopic viewpoint, the schedule server balances the load of every distributed server to avoid overload. From a microscopic viewpoint, optimization algorithm is used to design the most reasonable schedule scheme between distributed servers and users. In this way, the performance of the system will be near to the best. The pre-alarming mechanism is designed for exceptions. The task inspector can watch the working state of the system. When exceptions happen, the urgent process program will be automatically triggered to protect the system.

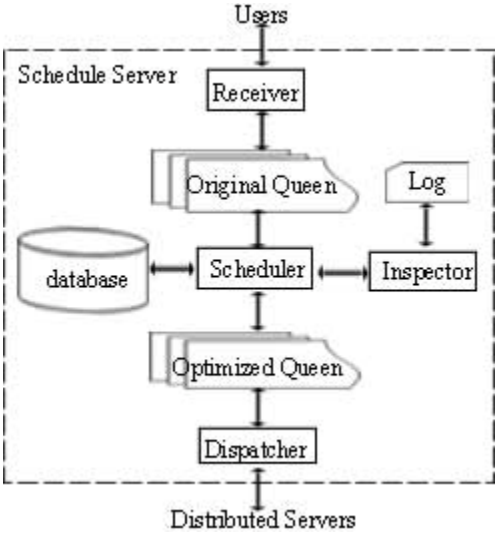


Fig. 2. Framework of schedule server

3.2 Description of Schedule

Assumptions: The number of distributed servers is *m* and they are $s_1, s_2, \cdots, s_m$. The number of users is *n* and they are $u_1, u_2, \cdots, u_n$. The maximum of server s_j connected to users is Max_{sj}. The maximum of user u_i connected to servers is Max_{ui}. M_{ui} represents task of user u_i. SM_{ui} represents task's size and TM_{ui} represents task's download time. V_{ui-sj} represents download velocity between user u_i and server s_j.

The optimization of the schedule emphasizes on the choice that how task M_{ui} selects the distributed servers? The constraint conditions are Max_{sj} and Max_{ui}. The target of optimization is to make the download time of all tasks as short as possible.

The function of optimization is as follows:

$$\min \quad E = \sum_{i=1}^{n} TM_{ui}$$

$$TM_{ui} = SM_{ui} \Big/ \sum_{j=1}^{m} V_{ui-sj} \cdot x_{ui-sj}$$

$$s.t. \quad \sum_{i=1}^{n} x_{ui-sj} \le Max_{sj} \quad (j = 1,2,\cdots,m)$$

$$\sum_{j=1}^{m} x_{ui-sj} \le Max_{ui} \quad (i = 1,2,\cdots,n)$$

$$x_{ui-sj} = \begin{cases} 1 & if\ user\ u_i\ downloads\ from\ server\ s_j \\ 0 & otherwise \end{cases}$$

This optimization is a typical MOP. In a MOP system, different targets always conflict. That is to say, the optimization of a target function often affects some other target functions [4].

3.3 Implementation of Algorithm

The traditional solutions to MOP are that it is transformed to many single-objective optimization problems. These methods strongly depend on cognition of the problem. Evolution calculation is an optimization technology based on swarm. It searches all solutions in the solution-space synchronously and improves the searching efficiency using the similarity of all different solutions. Therefore, Evolution calculation is a suitable algorithm for MOP. ACO is a typical evolution algorithm and we use it to solve the schedule between servers and users [5].

The key step of ACO is to transform actual problem to ant colony networks. We divide the schedule into *m* phases. In each phase, one server is assigned to supply download service. Max_{sj} ants are placed in server . Every ant moves from server s_j to users $u_1, u_2, \cdots, u_n$. Ant *k* chooses to go user with probability p_{ij}.

$$P_{ij}(t) = \begin{cases} \dfrac{[\tau_{ij}(t)]^{\alpha} \cdot [\eta_{ij}(t)]^{\beta}}{\sum_{s \in allowed_k}^{n} [\tau_{is}(t)]^{\alpha} \cdot [\eta_{is}(t)]^{\beta}} & j \in allowed_k \\ 0 & otherwise \end{cases} \tag{2}$$

Where, $allowed_k$ represents a list of the users which can be accessed by ant *k*. Visited users will be eliminated from the list during each phase. $\tau_{ij}(t)$ represents the pheromone density between u_i and s_j at the certain moment *t*. η_{ij} represents the heuristic degree of moving from s_j to u_i. α represents the weight of pheromone density and β represents the weight of heuristic degree. At the beginning, the pheromone densities of all paths are equal and $\tau_{ij}(0) = C$ (*C* is a constant). η_{ij} is determined by some heuristic algorithm. In this paper, η_{ij} is as follows:

$$\eta_{ij} = 1 / TM_{ui} \tag{3}$$

The deposit of pheromone:

$$\tau_{ij}(t + \Delta t) = (1 - \rho) \cdot \tau_{ij}(t) + Q \Big/ \sum_{i=1}^{n} TM_{ui}, \rho \in [0,1] \tag{4}$$

Where, ρ represents evaporation degree of pheromone. *Q* is a constant.

Steps of the algorithm:

(1) Initialize the value of ρ, Q, C, α, β.

(2) *oc*←*oc* represents outside counter).

(3) $nc \leftarrow nc$ represents inside counter).

(4) Place Max_{sj} ants at server s_j. Each ant chooses to go next user with P_{ij}. The limits are ants in the same server cannot move to the same user and the number of ants in one user cannot exceed Max_{ui}. Then modify the list $allowed_k$.

(5) Calculate the value of optimization function *E* according to formula (1), and put the best result into the result list after comparison. Then *ncnc*+1.

(6) If *nc*>the initial value, then use the better result to update the pheromone according to formula (4), and *ococ*+1; else reset the list $allowed_k$ and return step (4).

(7) If *oc*> the initial value, then display the best solution; else return step (3).

4 Experiment

To evaluate the efficiency and validity of schedule algorithm, we have done a competitive experiment. The data are as follows:

m=3, n=20, Max_{s1}=Max_{s2}=6, Max_{s3}=8, all users' tasks are the same and the size of tasks is 50M.

The maximum of all users connected to servers is 1. The download velocities between users and servers are limited from 100k/s to 900k/s and randomized. In addition, constants in ACO are initialized as follows:

α =2, β =2, ρ =0.1, Q=1, C=0.1. Through 200 generations, we got the values of optimization function *E* as Fig. 3.

From Fig. 3, it shows the better result is obtained at the 106th generation. The value of optimization function *E* is 1243.5s. In order to prove the optimal performance of ACO, we compared with traditional load-balance algorithms in dealing with the schedule between distributed servers and users, such as cycle, random, the least response time and cycle with weight [6]. The comparison result of ACO and traditional algorithms is showed in Fig. 4.

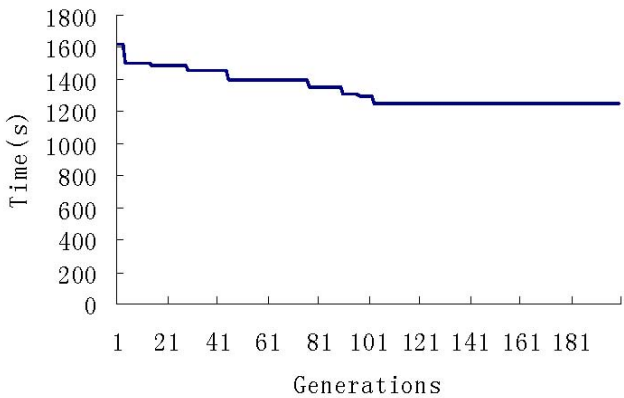


Fig. 3. Values of optimization function *E*

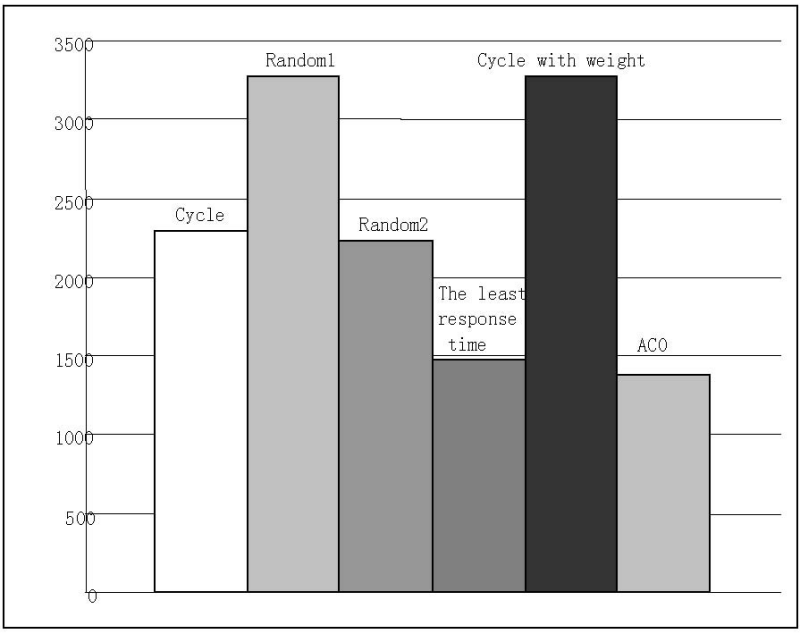


Fig. 4. Comparison Result of ACO and traditional algorithms

From Fig. 4, it shows schedule based on ACO uses fewer time than FIFO obviously, not only in the download time of all tasks but also in solitude download time of many tasks.

5 Conclusions and Future Work

The DPSLR effectively solved the schedule between distributed servers and users, and made learning resources update in time. A new Internet-based learning resources update service system is formed. It will bring some application values as follows:

(1) Effectively solve the shortages of users' learning resources and supply new learning resources for users continuously.
(2) Shorten the update cycle of learning resources to guarantee its temporality.
(3) Economize the cost of learning resources update service.

Due to the uncertainty of networks, we will improve the schedule method in accordance with actual situations and enhance system performance. Technologies such as agent, data mining and rule-based reasoning will be used to analyze users' interests and to form user model. It is an intelligent learning resource update service system which can supply personalized order service, on-line service and active push service.

Acknowledgment

Our research is supported the National Torch Key Program Foundation of China.

References

1. Seels, B.B., Bichy, R.: Instructional Technology: The Definition and Domains of the Field. In: Association for Educational Communications and Technology, Washington, DC (1994)
2. Pei-jun, D., Dong-dai, Z., Xiao-chun, C.: The Research on Individual Service Model of Education Resources Based on Multi-Agent Technology. Journal of Northeast Normal University(Natural Science Edition) 38, 31–35 (2006)
3. Yi-xing, Y., Yang, G.: Research on Resource Scheduling in the Grid. Application Research of Computers 5, 23–26 (2005)
4. Abido, M.A.: Multiobjective Evolutionary Algorithms for Electric Power Dispatch Problem. IEEE Transactions on Evolutionary Computation 10, 315–329 (2006)
5. Dorigo, M., Stutzle, T.: Ant Colony Optimization. MIT Press, Cambridge (2004)
6. Dewei, P., Yanxiang, H.: Study of Application on Framework of Load Balancing Based on Agent. Computer Engineering and Applications 5, 153–156 (2005)

On Retrieval of Flash Animations Based on Visual Features

Xiangzeng Meng and Lei Liu

School of Communications, Shandong Normal University
250014 Jinan, China
mxz@sdnu.edu.cn

Abstract. Flash is undergoing an explosive spread as a new prevailing media format on the Web. Unfortunately, few research efforts have been devoted to content-based Flash retrieval (CBFR) in IR community, which goes against the utilization of the enormous Flash resources. In this paper, Flash animation in 3-layer architecture is presented after it is segmented based on its visual features to a series of scenes on time-line of the production process. A promising approach of Flash animation retrieval is proposed based on the visual features of scenes and some meta-parameters. An experimental prototype system of Flash animation retrieval with roughly 100,000 Flash animations in total is built. The primary experiment demonstrates the flexibility and the effectiveness of our approach of CBFR.

Keywords: Flash animation; CBFR; visual feature; scene segmentation.

1 Introduction

Flash, as a vector-based media format, is widely used for cartoon, music TV, game, VR on desktop, commercial advertisement, e-postcard, etc, with the advantages of small size, easy composition, bright dynamic effect, powerful interactivity, etc[1]. Since its advent in 1997, Flash animation has experienced an explosive growth and become one of the most prevalent formats on the Web. The retrieval, management and utilization of Flash resources become the interesting issues in IR community. Nowadays, some multimedia search engines offer Flash retrieval on keywords. However, it is difficult to describe its contents and visual features using two or three keywords due to the complicated content and the visual effect of Flash animation. Content-based multimedia retrieval (CBMR) has been developed since 1990s [2]. Unfortunately, the research efforts on CBMR are mainly aiming at the retrieval of image, video and audio, and there are quite limited research papers involved in content-based retrieval and management of Flash resources. Yang et al. [3] proposed a generic framework named FLAME (Flash animation Access and Management Environment) through the analysis of content structure, which embodies a three-tiered architecture for the retrieval of Flash animations. The FLAME has implemented the retrieval at different levels of details, including object level, event level and interaction level. Ding et al. [4] suggested a semantic model with co-occurrence analysis for improving the performance of retrieval of Flash. No other relevant research on

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 270–277, 2008.

content-based index and retrieval of Flash animation has been reported (to the best of our knowledge).

Flash animation is a kind of streaming media with heterogeneous components (including texts, graphics, images, videos and sounds), dynamic effects and interactions. On the view of movie playing, a Flash animation is composed of a series of scenes like a video stream, therefore it may be presented and indexed similar to video based on its shots and scenes. The approaches proposed by [3], [4] have described comprehensively the content structure of Flash animation overlooking the visual effects and scene structure. On the view of visual effects of Flash animation, Flash animation is segmented into a series of scenes based on the visual effects showing on screen, and the key visual features of scenes are extracted to represent Flash animation based on its scene structure. A Gif animation image is constructed with the typical frames (one extracted from every scene) to be used as a summary displaying the main playing screen of Flash animation. Finally, a system of Flash animation retrieval and quick browsing is realized, based on scene structure and visual features.

The rest of the paper is organized as follows. In Section 2, the structure of scenes of a Flash animation is analyzed and a 3-layer architecture to represent the content of Flash animation is presented. The method of keyframe extraction and scene segmentation of a Flash animation is elaborated in Section 3. The retrieval of Flash animations based on visual features of scenes is described in Section 4. The conclusion is given and promising future directions are suggested in Section 5.

2 The Structure of Scenes and Representation of Flash Animations

Flash animation is a movie which is played continuously on the time-line from one frame to another. It includes two kinds of frames: keyframe and generated frame. The

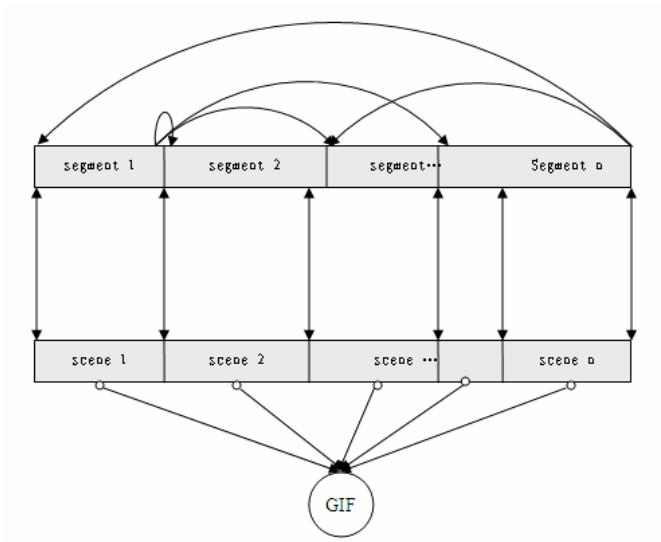


Fig. 1. The architecture of scenes in a Flash animation, ○ means the typical frames

former is the frame in which some objects, object properties, changes, actions or interactions are set or defined by the producers. The later is generated automatically through interpolation calculation with the help of the before and after keyframes by Flash animation making tools. A sequence of frames which can be automatically played in Flash animation is called a segment. All segments are arranged on the time-line of production and they are linked through the jump scripts in keyframes on hyperlink structure of Flash animation, as illustrated in Fig.1 (upper part). The hyperlink structure of Flash animation indicates the relationship between segments, but it cannot reveal the visual effect of scenes in Flash animation. In a Flash animation, the visual features may sometimes vary greatly in a segment, while sometimes they may vary less among several neighboring segments. On the view of visual effects, the frame sequence on the production time-line is segmented to a series of scenes, each of which is composed of the contiguous frame sequences with similar visual features (looking similar on visual effects). Thus a Flash animation can be viewed as a movie composed of a series of scenes with different visual features, as illustrated in Fig. 1 (lower part).

On the basis of scene segmentation and scene analysis, combined with extraction of meta-information of Flash animation, the content of Flash animation can be described in a 3-layer structure, as Fig. 2.

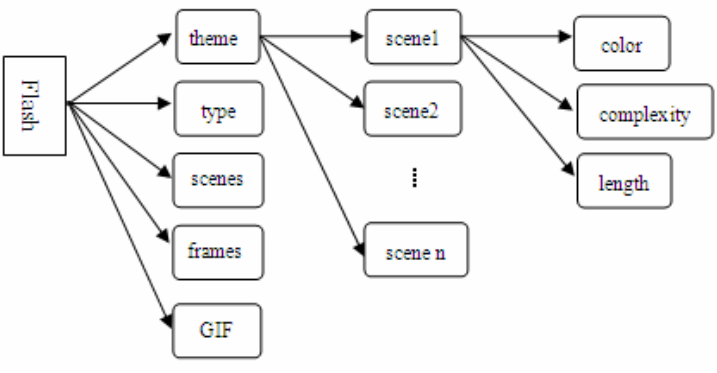


Fig. 2. The presentation of Flash animation based on scenes

The first layer is the global features of Flash animation. The thematic contents of Flash animation are represented by some key phrases, which are extracted from the internal embedded texts of Flash animation, such as nouns, verbs, adjectives, phrases, etc. The type refers to the types of subject matter, including cartoon□music TV, game, VR in desktop, CAI courseware, etc. GIF is composed of the representative frames of all the scenes, used to display the major content and general picture of Flash animation scenes. A representative frame is the keyframe in the middle of scene, basically representing the visual features of scene picture. The second layer refers to the scene structure of Flash animation. The third layer is the visual features of scenes. The color is the average color of the representative frame image. The complexity is measured by the average edge density (the greater the density is, the more complicated the scene is, and the greater the complexity is.), and the length refers to the frame count of scenes.

3 Keyframe Extraction and Scene Segmentation

A keyframe is the frame in which a key action or a change on some content is defined. But in the SWF file, the keyframes and the generated frames use the same format of definition and the same method of storage. No special markings to distinguish the keyframes from the generated frames in SWF file. In order to find out the keyframes in all frames, the raw data of SWF file need to be parsed.

The SWF file is makeup by a series of tagged data blocks and the structure is similar to the one of a XML file. In essence, a SWF file can be regarded as an encoded XML. There are three types of tags, namely Definition Tags, Control Tags and Display List Tags. Definition Tags are used to define the content of the SWF file, such as the shapes, texts, bitmaps, sounds, and so on. Control Tags are used to create and manipulate rendered instances of characters in the dictionary, and control the flow of SWF file. Display List Tags are used to add characters and character attributes to a display list.

A keyframe should meet one of the following five circumstances: adding characters to the display list, removing characters from the display list, modifying attributes of characters, morphing or including ActionScript code. Adding characters, removing characters may be simply judged respectively by PlaceObject tag, RemoveObject tag and RemoveObject2 tag. Modifying the attributes of characters in the display list may be judged by PlaceObject2 tag and PlaceObject3 tag. The morphing and ActionScript code can be simply recognized by DefineMorph tag and DefineAction tag.

In the progress of keyframe extraction, it is very important to detect the motion tweens. It is hard to distinguish keyframes from general frames if only using Display List tags, because in the motion tweens, PlaceObject2 tags and PlaceObject3 tags are used to modify the character's properties in both keyframes and general frames. According to the SWF file structure, the motion tween are recognized by finding that some same tags are used to modify a character's properties continuously. In such situation, the first frame and the last frame are the keyframes.

Due to the complexity of internal structure and storage structure, the screenshot function provided by Flash animation Player ActiveX is used to save the keyframes instead of analyzing of the contents of the keyframes after the keyframes are indexed. Experimental results show that the screenshot method has a higher efficiency than the content analysis of keyframes does.

Scene segmentation is mainly dependent on the visual features of the scene. Color is an important perceptual feature of the image and it is not sensitive to the change of position and direction. Color features of an image are more robust than geometric features. Scene segmentation is fulfilled by comparing the color difference of two adjacent keyframes rather than frame by frame considering the little difference among the

1	2	1
2	8	2
1	2	1

Fig. 3. The sub-blocks and their weights of a frame image

middle frames between two keyframes and improving the speed of calculation. The color difference of two keyframes are calculated on 9 blocks divided with different size and weight, as illustrated in Fig.3. The color difference between two keyframes is the sum of the weighted difference of each block.

In order to calculate the color difference rationally, an improved HSI color model is adopted. In the model, the hue (H) is kept invariant and the saturation (S) and brightness (I) are adjusted. And the entire color space is compressed into a sphere with the radius of 0.5. The transformation between our improved HSI model and RGB model are as follows.

$$H = \begin{cases} \theta & G \geq B \\ 2\pi - \theta & G < B \end{cases} \tag{1}$$

$$S = \begin{cases} S_0 \sqrt{0.5^2 - (Y_{m0} - 0.5)^2} \sqrt{1 - (\frac{I}{Y_{m0}} - 1)^2} & I < Y_{m0} \\ S_0 \sqrt{0.5^2 - (Y_{m0} - 0.5)^2} \sqrt{1 - (\frac{I - Y_{m0}}{1 - Y_{m0}})^2} & I \geq Y_{m0} \end{cases} \tag{2}$$

$$I = \frac{Y^2}{Y_m} \tag{3}$$

$$\theta = \arccos\left(\frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2 + (R-B)(G-B)}} \right) \tag{4}$$

$$Y = 0.30R + 0.59G + 0.11B \tag{5}$$

$$S_0 = 1 - \frac{3\min(R,G,B)}{(R+G+B)} \tag{6}$$

Where, Y_m is the maximum possible value when hue and saturation are certain values. Y_{m0} is the value of Y_m when the saturation is 1.

In the improved HSI color space, the distance (d_{12}) of any two colors (C_1, C_2) is between 0 and 1. (the distance between white and black or any other two complementary colors is 1)

$$d_{12} = \sqrt{(I_1 - I_2)^2 + S_1^2 + S_2^2 - 2S_1S_2\cos(H_1 - H_2)} \tag{7}$$

When calculated the value of color difference between two keyframes, a threshold value is set to determine the border of the scene. If the value of color difference is greater than the threshold value, the two keyframes divided to different scene. Otherwise, the two keyframes belong to the same scene. Then the middle keyframe as the representative frame of the scene is selected and the mean color and image complexity are calculated. Finally, all of representative frames arranged on the production time-line are converted to a GIF animation to be used to browse the scenes shown in Flash animation.

Experiments show that scene segmentation based on weighted region has higher sensitivity and the accuracy rate is 85% compared with the artificial segmentation.

4 Retrieval of Flash Animations Based on Visual Features of Scenes

Based on the representation of the visual features of Flash animation in figure 2, the method of Boolean matching combined with fuzzy-comparison is adopted to retrieve Flash animations. The size of file and the type of Flash animation adopt Boolean matching and the theme, the color, the complexity adopt fuzzy-comparison. The calculating formulas of similarity between the desired and the indexed Flash animations are as follows.

$$S = D_1 \times D_2 \times S_t \times \frac{(S_c + S_f)}{2} \tag{8}$$

$$S_t = \frac{|\mathbf{A} \cap \mathbf{B}|}{|\mathbf{A}| + |\mathbf{B}|} \tag{9}$$

$$S_c = 1 - \max_i \left\{ \min_j \left\{ d\left(c_i, c_j\right) \right\} \right\} \tag{10}$$

$$S_f = \text{mean}\left\{ R_j \right\} \tag{11}$$

$$R_j = \frac{1}{1 + \left(\frac{\delta_m - \delta_j}{\Delta} \right)^2} \tag{12}$$

$$\Delta = \begin{cases} \delta_j & ; \delta_j \le \delta_m \\ 1 - \delta_j & ; \delta_j > \delta_m \end{cases} \tag{13}$$

Where, D_1, D_2 is the logical value which expresses the match degree of the size and the type respectively. S_t, S_c, S_f is the fuzzy similarity of theme, color and complexity respectively. $|\mathbf{A}|$, $|\mathbf{B}|$, $|\mathbf{A} \cap \mathbf{B}|$ is the number of query keywords, the number of keywords indexed and the number of the synonymous words of them respectively. c_i, c_j is the desired and the indexed color respectively. $d(c_i,c_j)$ is the Euclidean distance of them in our improved HSI color space. R_j is the satisfactory calculated with formula (12) at the complexity δ_j of scene j under δ_m (δ_m=0, 0.5, 1 meaning respectively to retrieve complex, medium or simple scene). S_f is the average value of R_j.

To demonstrate the feasibility and effectiveness of our approach, an experimental prototype as a Flash animation retrieval system based on Visual Features of Scenes has been built. The prototype system has indexed roughly 100,000 Flash animations downloaded from related Web pages in total. The interface of the system is displayed in standard Web browser and may be accessed remotely over the Internet supporting query by keywords (to theme), color and complexity (to scenes). 50 Flash animations with top similarity calculated based on formulation (8) are retrieved and ranked in

descending order. The GIF animation of scenes, the address of related web page, title (linking to the file address) and the file size of each animation retrieved are displayed 10 per page, as illustrated in Fig. 4. The average retrieval accuracy of 300 retrieving trials with the combination of 20 themes, 5 colors and 3 complexities, is 83%, much higher than 57% only with 20 themes.

Fig. 4. The retrieved Flash animations in a retrieval example

5 Conclusions

This paper has investigated the problem of retrieval of Flash animation based on visual features, which is helpful to better utilization of the proliferating Flash animation resource in web. As the major contribution of this paper, a promising approach of Flash animation retrieval based on the visual features of scenes is proposed. The primary experiment on a database with roughly 100,000 Flash animations demonstrates the flexibility and the effectiveness of our approach. However, there still remain many issues for future research, for example, extraction of visual features of components, presentation of logical architecture of the segments and their scenes and components, retrieval based on integration of visual features, components, logical architecture and semantic of Flash animations, etc. Some of the issues are our research direction in the future.

Acknowledgments. The work described in this paper was substantially supported by a grant from Natural Science Foundation of Shandong Province. (Project No. Y2005G21)

References

1. Flash Player statistics, http://www.adobe.com/products/player_census/flashplayer/
2. Lee, T., Sheng, L., Bozkaya, T., Ozsoyoglu, G., Ozsoyoglu, M.: Querying Multimedia Presentations based on Content. IEEE Trans. Knowledge and Data Engineering 11(3), 361–387 (1999)

3. Yang, J., Li, Q., Wenyin, L., Zhuang, Y.: FLAME: A Generic Framework for Content-based Flash Retrieval. In: 4th International Workshop on Multimedia Information Retrieval, in conjunction with ACM Multimedia 2002, Juan-les-Pins, France (2002)
4. Ding, D., Li, Q., Feng, B., Wenyin, L.: A Semantic Model for Flash Retrieval Using Co-occurrence Analysis. In: Proc. ACM Multimedia 2003, Berkeley, CA (2003)
5. Chang, S.F., Chen, W., Meng, H.J., Sundaram, H., Zhong, D.: VideoQ: An Automated Content based Video Search System Using Visual Cues. In: Proc. ACM Int. Multimedia Conf., pp. 313–324 (1997)
6. Flash Player Developer SDKs, http://www.adobe.com/licensing/developer/
7. Meng, X., Wang, L., Li, H., Zhong, Y.: A Method of Image Retrieval with Color Names. Chinese Graphic and Image 10(3), 349–353 (2005)

The Design of Web-Based Intelligent Item Bank

Shaochun Zhong[1,2,3,4], Yongjiang Zhong[2,3,4], Jinan Li[2,3,4],
Wei Wang[2,3,4], and Chunhong Zhang[5]

[1] Software School in NorthEast Normal University (NENU), JiLin ChangChun, 130024
[2] Ideal Institute of Information and Technology in NENU, JiLin ChangChun, 130024
[3] E-learning laboratory of Jilin Province, JiLin ChangChun, 130024
[4] Engineering & Research Center of E-learning, JiLin ChangChun, 130024
[5] ArmorTechnique Institute , JiLin ChangChun, 130117
sczhong@sina.com, zhongyj@nenu.edu.cn,
ljn@szftedu.cn, wangw577@nenu.edu.cn,
zchji@sogou.com.

Abstract. Based on briefly analyzing the current situation and problems of internet item bank, the paper proposes a web-based intelligent item bank which is intelligent and can evaluate the tests effectively. Then it introduces the function design, architecture and implementation scheme of the system. It elaborates the evaluation strategies--S-P table in detail, which can not only evaluate the student individual and in whole but also evaluate the testing. Finally it discusses the future development of the system.

Keywords: Internet item bank, Paper appraisal, Adaptability.

1 Introduction

Item bank is "a set of items for a special subject constructed by computer system according to the education measurement theory". Item bank following the education measurement theory is an education measurement tool based on strict mathematical model.

Item bank construction is a complex system engineering. First, establish the mathematical model. Then set the attributes indexes of the items and the structure of the test papers, input the scientific effectual items. Following that, evaluate the testing effect and situation of the students' leaning. After putting into operation, adjust adaptively according to the testing level of the students. To ensure the scientificalness and effectiveness, the relative attributes of the items and test papers should not only be set by experts but also be sample tested using large numbers of sampling data, which can adjust the effectiveness of the parameters. Editing and testing task for a relatively complete item bank based on the classical test theory is such a formidable task, which can not be done well by average institutions. [1]

Presently with the development of information technology, item bank are established by companies, schools or their cooperation. They have made certain achievements in practical application, but with some problems.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 278–289, 2008.

(1) Independent of the other teaching links, used un-widely, so can not be improved based on sufficient sample data.
(2) Providing reverence answers without education evaluation and intelligent instructing.
(3) The adaptability of the system is not so good, for example: the difficulty index keeps invariable and papers can not be generated automatically.
(4) Having little characteristics of automatic marking and reasonable statistics analysis functions.

In order to solve the problems above, we improved the existed internet item bank and developed a general item bank system successfully. It consists of subsystems of item management, automatic test paper generation, autonomous learning, testing, and study evaluation, which can provide a series of services: automatic test paper generation, testing, analysis and evaluation.

2 System Design

2.1 Overall Design

In order to solve the problem of running independent of other teaching links, the system is designed with the relative systems such as testing, evaluation and autonomous learning. The overall design is illustrated by figure 1.

2.2 Function Analysis

(1) Item Management Subsystem

The sub-system includes the question database and the paper database. Question database is used for storing all kinds of questions, and paper database is used for the papers generated by paper generation subsystem and obtained by the paper input interface. This subsystem can edit the questions and papers such as multiple choice, calculation, close test, line connecting, etc, and can store the corresponding answers and explanations.

Fuction include:Searching,Adding,Editing,Deleting,Submission online.

(2) Paper Generation Subsystem

Paper can be generated automatically or manually. Manual generation means to generate the papers from the item bank according to the teachers' requirements. Automatic generation methods can be divided into three types.

a) Generating by teachers: The teachers input the parameters of paper generation (including: paper title, testing time, full mark, total number of the items, the knowledge nodes tested, average difficulty, etc) via browser, then the system can generate the paper and answer according to the requirements. There are two ways to show the paper. One is in the form of web pages, which can be edited online by teachers and then printed .The other is in the form of RTF file packet, which can be downloaded for use.

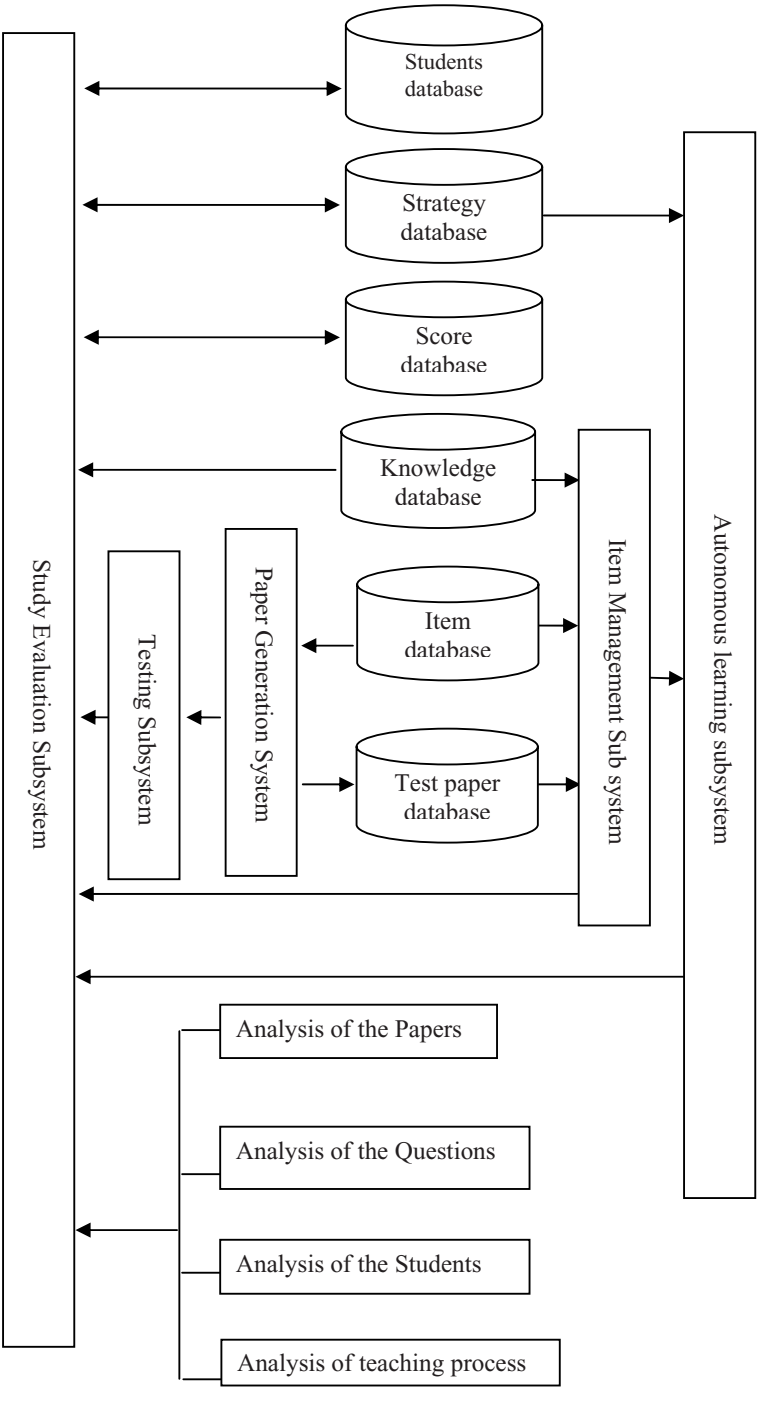


Fig. 1. Overall Design

b) Generating by students: The students can input the parameters of paper generation (different from the one used by teachers) according to their own condition, then the system generates the paper for exercise of pertinence which can be printed or saved as web pages.

c) Static generation strategies: There are some common generation strategies for the common testing. The teachers can choose the ready generation strategies instead of inputting the complicated parameters.

(3) Testing Subsystem

This subsystem can use the papers generated by the system or the ones generated by the paper generation subsystem. Students take the test online or take the paper home via storage device to test, and then submit their answers to the analysis modules. This subsystem includes three links.

a) On-line examination: The students can choose the special paper via browser, take the test and submit their answers. The system stores the answers into the student information database, and then the teachers can take them out and do the marking.

b) On-line marking: After login in the system, the teachers can choose the paper they should mark, and then choose the students who took the examination to take the students' papers and corresponding answers. The objective questions can be judged by the system automatically. The non-objective questions are marked by teachers which are stored into the score database with the marking information together after submitted by teachers.

c) Results requiring: After login in, the students can choose the papers they tested and look over the testing contents, standard answer and marking condition, etc.

(4) Autonomous learning subsystem

Autonomous learning subsystem has a close relationship with the testing subsystem. On one hand, it calls the paper generation subsystem to generate the papers, do exercise autonomously and transfer the results to the analysis modules. On the other hand, it can give the corresponding advices on knowledge review and learning strategy according the test.

(5) Study Evaluation Subsystem

This subsystem is used to do statistic analysis of questions, papers and the teaching process according to the history data of the students' testing which is analyzed by S-P table. It has two important parameters of difference coefficient and alarm coefficient, and they can separately provide important information for analysis of teaching situation and teaching effect. So, this subsystem can analyze the reasonable degree for individual situation, trends of all the students and improvement of the teaching effect. It includes the following aspects.

a) Statistics analysis of the papers: The distribution of the items such as students' points can be shown by line graph and histogram and all the abnormal questions in the exam can be captured such as all the answers are right or wrong in whole. Analysis on the following subjects: the reliability, validation and average difficulty of the paper; the maximum and minimum points, the total number of each grade section, average point and standard deviation ; the original points and the transformed points for each student.

b) Statistics analysis of the questions: Analysis results such as difficulty, differentiation and marking of knowledge points can be shown in tables and figures. Analyzing the marking of each question provides the current study situation of the student group. Then the students' recognition to knowledge points and problems can inferred even to affect the teaching strategies.

c) Statistics analysis of the students:

i. Score Clarifying: When the students choose different groups, their original points can be transformed from different sample ranges. So the students can understand their own position of the different groups.

ii. Aware of the score trends: Track the history of testing scores in time order , transform the original points to the standard points , 100-level graduation points, etc .After such transform, to compare the points of each subject is more reasonable and easy to track the score trends.

iii. Analysis of knowledge and ability: On-line examinations can be recorded by computer and use the record to analyze the significative response information. Paper examinations can be input the response for the paper manually. Analyze from two dimensionalities of cognitive ability and knowledge contents. Teaching object range, degree and ability condition of each knowledge point can be analyzed by its knowledge attribute, cognition classification attribute and the students' responses.

iv. Intelligent instructing: The analysis result of students' learning can be used to suggest the students what is need to enhance and to list the relative teaching materials , analysis of their own fault and weak knowledge points, such it's easy to teach students in accordance with their aptitude. [2]

d) Statistics analysis of the teaching process: Based on the score accumulation of many tests, it can analyze the point distribution of each knowledge point. If the students ' responses are abnormal, it suggests that problems exist in the knowledge unit and teaching should be improved.

3 Evaluation Strategies

S-P table is an information processing method of analyzing the teaching based on the students' points table of the questions. The method can evaluate the learning situation of individual student and the trends of all the students; also it can evaluate propriety of all the questions. Difference coefficient and alarm coefficient are two important parameters used for S-P analysis. Calculation of them can provide important information for analysis of teaching situation and teaching effect.

3.1 Formation of S-P Table

M students answer n questions, mark 1 when right, mark 0 when wrong, and a matrix of student – question is created .In the matrix, U_{ij} represents the point student i gets for answering question j. Because of adoption of 0-1 score method, the matrix is consisted of only 0 and 1.The score matrix can be shown concretely as the condition in figure 2.

Student-Problem Matrix

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Score
S1	0	1	1	0	1	1	0	1	1	0	6
S2	0	1	1	1	1	0	0	0	0	0	4
S3	1	1	1	1	1	0	1	1	0	1	8
S4	0	1	1	1	1	1	1	1	0	0	7
S5	1	1	1	1	1	1	1	0	0	1	8
S6	0	1	1	0	0	0	0	1	0	0	3
S7	1	1	1	1	1	1	1	1	0	1	9
S8	1	0	0	1	1	1	1	1	0	1	7
S9	1	1	1	1	1	1	1	1	1	1	10
S10	1	1	1	0	1	1	0	0	0	0	5
S11	0	1	0	1	1	1	1	0	0	1	6
S12	0	1	1	0	1	1	1	1	1	0	7
S13	0	1	1	1	1	1	1	0	1	1	8
S14	0	1	1	0	1	0	1	0	0	0	4
S15	0	1	1	1	1	1	1	1	1	0	8
Correct Times	6	14	13	10	14	11	11	9	5	7	

Fig. 2. Student-question matrix

The column on the right lists the scores of the students (the number of the right ones); the bottom lists the number of right answers for each question. Before the matrix is processed, it can only provide the point of each student and the correct rate of each question. It can provide much significant information after processed by the following rules.

a) Students will be arranged by descending order according to the scores. Namely, exchange the students scoring line, making the high score above in the row, low score firms bottom.

b) Problems will be arranged by descending order according to the number of correct answers from left to right. Namely, exchange the problems line, making the problem line of giving more correct answers to it on the left, the problem of low frequency of correct answers row on the right side.

c) For the same score line, first of all, find out the sum of correct answers which the student obtained the errors of each problem in the same column. And then, arrange the line of the same score according to the sum; line get high score firms the top.

d) For the same score column, first of all, find out the sum of the student obtained the errors of each problem. And then, arrange the column of the same score according to the sum; column get low score firms the top.
e) Making the S line. Draw the vertical line for each of the students. Making the number of the problems on the left side of the vertical equal to the students' score. And then, drawing the striping between lines, link each vertical line to form a ladder of the curve, called S line. As Figure 3 shows the solid line.
f) Making the P line. Draw the ledgement for each of problems. Making the number of students above the ledgement equal to the number of correct answers. And then, drawing the vertical line between columns, link each ledgement to form a ladder of the curve, called P line. As Figure 3 shows the dashed line.

After processing, the student-question matrix is transformed into ordered list of S (solid line) and P (dashed line), which is called S-P table.

S-P Table

	P5	P2	P3	P7	P6	P4	P8	P10	P1	P9	Score
S9	1	1	1	1	1	1	1	1	1	1	10
S7	1	1	1	1	1	1	1	1	1	0	9
S3	1	1	1	1	0	1	1	1	1	0	8
S13	1	1	1	1	1	1	0	1	0	1	8
S15	1	1	1	1	1	1	1	0	0	1	8
S5	1	1	1	1	1	1	0	1	1	0	8
S8	1	0	0	1	1	1	1	1	1	0	7
S12	1	1	1	1	1	0	1	0	0	1	7
S4	1	1	1	1	1	1	1	0	0	0	7
S1	1	1	1	0	1	0	1	0	0	1	6
S11	1	1	0	1	1	1	0	1	0	0	6
S10	1	1	1	0	1	0	0	0	1	0	5
S2	1	1	1	0	0	1	0	0	0	1	4
S14	1	1	1	1	0	0	0	0	0	0	4
S6	0	1	1	0	0	0	1	0	0	0	3
Correct Times	14	14	13	11	11	10	9	7	6	5	

Fig. 3. S-P table

3.2 Properties of S-P Table

Based on the data relationship, it is not difficult to see that the S-P table has the following basic properties.

Because the score of students equals the total time of the right answers, the left-hand area of S is equal to the area above the P; S is not only the score curve of the students but also the number curve according to the cumulative scores. P is the accumulating distribution model of the right answer number for each question. S and P are always intersectant, and when superpositioned, the left end of P is always above S and the right end of P is always under S; the area gap between S and P is called dispersion of the two lines which can describe the relationship between the question difficulty and the students' response. If S is as same as P, the condition is that the learning state of all the students is absolute stable.

3.3 Analysis of S-P Table

According to S-P table, holistic analysis can only be done for whole and individual analysis can only be done for individual.

(1) Holistic Analysis

It includes the difference analysis between the students and the questions and the distribution analysis between the questions and students' response. The dispersion of the two lines reflects the evenness degree of distribution. The increase in dispersion can be associated with the increase in un-even distribution. The evenness of distribution can be described as follows:

$$\text{D(magnitude of deviation)}=\frac{\text{area between S-line and P-line}}{(S-P)\text{surface area}} \tag{1}$$

$$(S-P)\textit{surface}\ \text{area=the number of students}\times\text{the number of problems} \tag{2}$$

As for common tests, the value of D ranges from 0.25 to 0.35, and can not over 0.5 usually. If it is over 0.5, the relationship between the questions and students' response is abnormal .For example, if a difficult problem is solved by students getting low points and not solved by students getting high points.

The difference between students and questions can be reflected by the faultage between S and P. Faultage stands for the beeline range between the domestic wirings. Figure 4 illustrates the faultage.

If the line S keeps horizontal for quite long, it suggests the danger of polarization. If the line P keeps vertical for quite long, it suggests that there is sufficient difficulty difference which may influence the validity of the testing.

(2) Individual Analysis

It evaluates both the learning situation of the students and the pertinence of the questions. When the dispersion between S and P is quite large, some responses of the students or the questions are abnormal. As the S-P table in figure 3, the rate of right answer of Problem P1 is 8/15, which is quite high. But the number who gives a right answer is just a half both above and below line P, which shows whether the problem

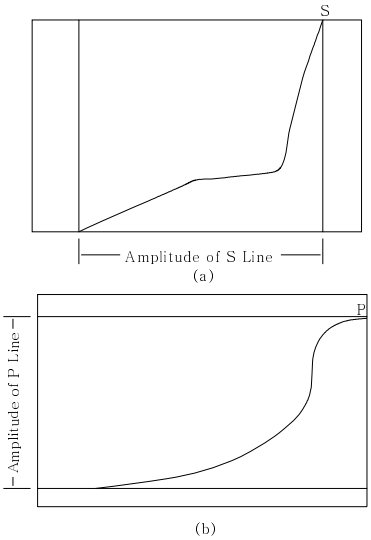


Fig. 4. Faultage

can be answered correctly is random distributed, and the students no matter with high or low points answered the questions correctly random. So the differentiation index of the problem is quite low and need to be checked. For example, there are two students who both answered 5 problems right, but the problems for each student distributes nearly halfly on both sides of line S ,which is also quite random and need to pay attentions. The abnormality degree of the students and questions can be described by alarm coefficient.

Calculation formula for the alarm coefficient:

$$W_j(P) = 1 - \frac{COV_j(P)}{COV_j(C)} \tag{3}$$

$W_j(P)$ represents the alarm coefficient of problem j; $COV_j(P)$ represents the covariance of the response pattern of problem j and the general response pattern; $COV_j(C)$ represents the covariance of the general response pattern and the complete response pattern. Complete response pattern means the response pattern that above the curve of S are all 1s and the blow are all 0s.

Calculation formula for the covariance of the response pattern of problem j and the general response pattern:

$$COV_j(p) = \frac{1}{m}\sum_{i=1}^{m}(X_{ij} - \frac{1}{m}PX_i)(SX_i - v) \tag{4}$$

In the formula, v means the average score of the students, and it can be calculated as follows:

$$v = \frac{1}{m}\sum_{i=1}^{m} SX_i \tag{5}$$

Then, $COV_j(P)$ is transformed:

$$COV_j(P) = \frac{1}{m}(\sum_{i=1}^{m} X_{ij} SX_i - PX_j V) \tag{6}$$

Calculation formula for the covariance of the general response pattern and the complete response pattern:

$$COV_j(C) = \frac{1}{m}(\sum_{i=1}^{m} X_{ij} SX_i - PX_j V) \tag{7}$$

Then, the alarm coefficient of the problem j can be calculated:

$$W_j(P) = 1 - \frac{COV_j(P)}{COV_j(C)} = 1 - \frac{\sum_{i=1}^{m} X_{ij} SX_i - PX_j V}{\sum_{i=1}^{PX_j} SX_i - PX_j V} \tag{8}$$

Calculation formula for the alarm coefficient of the student:

$$W_i(S) = 1 - \frac{COV_i(S)}{COV_i(C)} \tag{9}$$

In the formula, $COV_i(S)$ presents the covariance of the response pattern and the general response pattern; $COV_i(C)$ represents the covariance of the general response pattern and the response pattern of the student. Response pattern here means that the left-side of S are all 1s and the right-side are all 0s. As calculating the alarm coefficient for problems, calculate $COV_i(S)$ and $COV_i(C)$, and then $W_i(s)$:

$$W_i(S) = 1 - \frac{\sum_{j=1}^{n} X_{ij} PX_j - SX_i u}{\sum_{j=1}^{SX_i} PX_j - SX_i u} \tag{10}$$

In the formula, u means the average correct time for each question:

$$u = \frac{1}{n}\sum_{j=1}^{n} PX_j \tag{11}$$

When the alarm coefficient of the student is high, it suggests that this student makes mistakes in the relative easy problems but solves the relative hard ones. The phenomena may result from the problems in the learning conditions .For example, students are not of high learning enthusiasm such as to answer careless, they don't exert the potential abilities or answer just by guessing, etc.

When the alarm coefficient of the student is over high, it suggests that this problem is answered wrong by students with high points but answered correctly by students with low points. So the differentiation index of the problem is low and its use value decreases in terms of score levels.

Experiments indicates that: when the value of $W_j(P)$or $W_i(S)$ is over 0.6,teacher should pay sufficient attentions to the problems of the students and courage them to make great effort. Also the problem should be deleted or modified. [3]

4 Implementation of the System

The software architecture of the item bank adopts the B/S network computing pattern with 3-layer architecture of express-tier, the business-tier and the data-tier.

The express-tier applies the asp technology and employs at least IIS 4.0 as web server. The business-tier is enveloped with COM components so that it can combine with the windows operating system and run steady and effectively. The data-tier uses the large-scale commercial database SQL Server, which has the characteristics of ensuring the data integrity and data security and makes for shortening the inputting and outputting time of the massing data.

The system uses XML for Sharing Information Resources within the item bank. It can transform according to different educational resource metadata information, such it's easy to share the resource information among item banks. The system also refers to specifications and technologies such as JavaScript, DOM, JavaMail, LDAP, Dict Protocol. [4]

5 Conclusion

At present, the system has already established and been on the test-run stage with the basic functions. It can solve the shortcomings of running independent of the other teaching links, lacking of education evaluation, intelligent instructing and statistics analysis well. Yet further research and development should be made.

(1) Further Application of Information Push Technology
Although the system push information such as evaluation, knowledge points to review after the functions such as statistics analysis and intelligent instructing, yet the technology is necessary to extend its application in other aspects. There are some examples. Push the teaching strategies after mining personalized information such as students' appetite on study media and learning regulation; Push more authoritive tests; Push the adaptive tests to students according to their knowledge regarding knowledge points. [5]

(2) Componentizing the System
The system provides a series of services of automatic test paper generation, testing, analysis and evaluation, which can provide strong support to the distant education. Each link of this system is also an important link in the current web-based distant education. The system can be separated into functional subsystems based on component, which can integrated into the distant education supporting platform seamlessly and combine with the network courses closely.

(3) Applying the Item Response Theory to the Adaptive Testing
Compare the IRT with classical test theory, the theoretical system of the former is built based on more complicate mathematical model with more accurate concepts and theoretical derivation. It's suitable for the adaptive testing. The testing system asks the student one or more questions, then chooses the questions which can evaluate the student most well and truly by analyzing the completed questions. In this way, on one hand it can make the test corresponds with the student's ability level and then provides the more accurate information to evaluate the students' ability. On the other hand, it can shorten the testing duration effectively.

References

1. Yu, S., He, K.: Design and Realization of the Network Test Bank System. J. China Distance Education 164, 53–57 (2000)
2. Jing, Y.J., Zhong, S.C., Li, X., Li, J.-N., Cheng, X.C.: Using instruction strategy for a Web-based Intelligent Tutoring System. In: Technologies For E-learning and digital entertainment, vol. 3942, pp. 132–139 (2006)
3. Fu, D.: Education Information Processing. Peking (2001)
4. Zhong, Y.J., Liu, J., Zhong, S.C., Zhang, Y.M., Cheng, X.C.: Programming of informatized instructional design platform for physics. In: Technologies for E-learning and Digital Entertainment, vol. 3942, pp. 171–177 (2006)
5. Liu, L., Wang, W.: Design of Web-Based Exam-question with Self-study and Adaptive Adjusting. J. Computer System Application 4, 45–47 (2006)

Methods on Educational Resource Development and Application

Shaochun Zhong[1,2,3,4], Jinan Li[2,3,4], Zhuo Zhang[2,3,4], Yongjiang Zhong[2,3,4], and Jianxin Shang[2,3,4]

[1] Software School in NorthEast Normal University (NENU), JiLin ChangChun, 130024
[2] Ideal Institute of Information and Technology in NENU, JiLin ChangChun, 130024
[3] E-learning laboratory of Jilin Province, JiLin ChangChun, 130024
[4] Engineering & Research Center of E-learning, JiLin ChangChun, 130024
sczhong@sina.com, ljn@szftedu.cn,
zhangzhuo_ca@sina.com, zhongyj@nenu.edu.cn,
shangjx576@nenu.edu.cn.

Abstract. By analyzing systematically the current state and existing problems of educational resource development and application (ERDA) both here and abroad, this paper presents methods of resource development and application based on the course of curriculum implementation, the curriculum resource structure and integration model and the evaluation factors of resource and curriculum integration. At last, the paper presents the methods on integrating intelligent resource and establishing distributed resource system by experimental study.

Keywords: curriculum resource, resource system, resource development method, resource evaluation factor.

1 Current State and Problems

Informatization has become the key point of each country's development. Obviously, educational informatization also becomes the key point of educational development [1]. Thus, schools set up informational environment such as campus network, computer classroom, multi-function classroom, etc. However, equipments lost their values rapidly, which will fall into disuse in 3 to 5 years. In the first year, the proportion of loss will be more than 30%. The equipment application in instructional system is the main criterion to evaluate education informatization level [2]. In order to utilize equipments maximally, the most important things are to develop good resource, corresponding application modes and approaches as well as service system for teachers and students utilizing them in the course of instruction [3].

In general speaking, on ERDA, there are lack of comprehensive arrangements, global design and systemic organization. On the theory and orientation of ERDA, there exist a number of deviations in the most of schools.

Concretely speaking, on resource development, there are no in-depth studies on resource development theory such as the instructional modes under the information

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 290–301, 2008.

environment, the instructional design basis and so on. Thus, the developed resource cannot bring advantages of network into play maximally, cannot solve key difficulties in learning and teaching process. The developed resource and software are too loose in global area and too centralized in local area, which have no share functions. Teachers cannot use them smoothly. On rules and approaches of subject-specific resource development, animation developing tools, resource integration tools and excellent resource structure, there are lack of normative studies. These lead to the lower efficiency and quality of resource establishment. On function, structure and technology of resource and their support software, less work has been done. The developed resource and software cannot meet teachers' needs. On categoricalness and consistency of resource, there exist more problems such as incomplete, incompatible and conflicting phenomenon [4].

On resource application and service system, there are no more modes and approaches for the integration of IT and subject, namely, no operating regulation. The existing application modes and approaches have no enough scientificaalness, hierarchy, excellent evaluation system (for example, what is the good integration class?), effective service support system and share system.

Instructional design under the environment of network means to provide students materials by utilizing network resource. The communication between teachers and students or among students mainly uses words. There are no communication ways for the process of solving problems. Instructional activities under the environment of network lack of support system on learning, directing, supervising, evaluating and feedback (for example, the supporting system in the course of investigative learning).

More important thing is that the development and change of IT and instructional modes are so fast that teachers cannot adapt to these dynamically. Clearly, how to provide teachers effective application support system has become more and more important [5].

According to the above, this paper presents methods of ERDA in the course of curriculum implementation, the curriculum resource structure and integration model, evaluation factors of resource and curriculum integration, and methods on integrating intelligent resource and establishing distributed resource system by experimental study.

2 Goals on Curriculum Resource Development

During the course of education informaziliation, questions that need to be answered are what kinds of resource should be established in each curriculum, why establish them like that and what the basis of resource establishment is.

To answer these questions, we should answer why apply IT into instruction? In fact, the purpose for doing so is not only to learn it but also to improve the quality and efficiency of instruction. To reach this purpose, first, we should find difficulties in the general instructional process, and then solve them by using IT. This is the reason for applying IT into instruction [8] [9].

In order to make clear what kind of resource should be established and why establish them like that, we must understand completely how to integrate IT and curriculum together, and in which aspects of curriculum IT can be used, etc.

2.1 Definition of IT and Curriculum Integration

Definition 1 IT and curriculum integration means to integrate IT with the aspects of curriculum completely for improving the instructional quality and efficiency. Aspects of curriculum include curriculum goal, curriculum content, curriculum design, curriculum implementation and curriculum evaluation. Figure 1 illustrates more details.

On curriculum goal, IT can enlarge the curriculum's goals, for example, writing compositions with the help of network and computer, it means to extend goals of literature education. On curriculum content, IT can change the medium and form of curriculum content display, and make curriculum content more rich and colorful. On curriculum design, IT can change fundamentally modes, approaches, steps and evaluation of curriculum design. On curriculum implementation, IT provide support platform for it, such as teaching support platform, self-learning and cooperating learning platform, communicating platform , feedback platform and so on. On curriculum evaluation, IT provide effective support means for its implementation, such as self-evaluation, students' evaluation and social evaluation, etc.

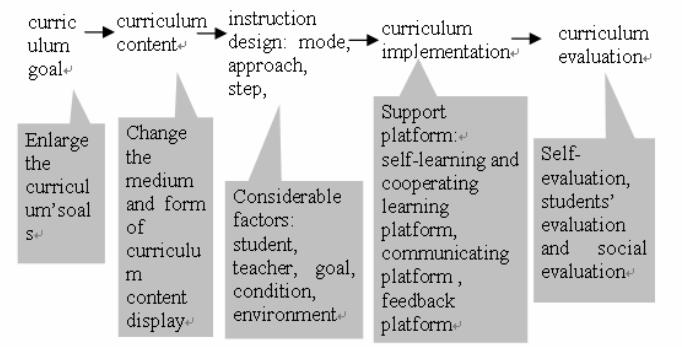


Fig. 1. Integration of information technology and curriculum

2.2 Goals on Curriculum Resource Development

Definition 2 curriculum resource means multimedia data and software saved in the computer, which can be spread through network for students' learning and teachers' teaching. These multimedia data and software can bear knowledge, transfer information, record data, make interaction, make control, make computation, make simulation and so on. Multimedia types include word, picture, image, video, audio, animation, software and their compound.

Definition 3 goals on curriculum resource development mean to develop resource that can help teachers and students to solve difficulties existing before class, in class and after class, which cannot be solved by the general instructional means. These resources are divided into the following types: just for teachers, just for students, for teachers and students, etc.

3 Methods of ERDA

ERDA should be considered from the sufficient integration of IT and curriculum. Developer should grasp the nature of integration. Methods of ERDA are multifarious, each school and teacher have their own methods according to their concrete conditions. Each method has its advantages and shortcomings. At present, there is no ideal method of ERDA yet. In the following, we contribute a set of ERDA methods based on the course of curriculum implementation:

3.1 Steps on ERDA Method

1. ERDA Method based on the course of curriculum implementation includes:

(1). curriculum diagnosis: it is to design instructional modes, arrange instructional activities, choose ways to complete instructional process and find difficulties that cannot be solved by general instructional means. In this process, instructional arrangement should be reasonable and optimized.

(2). exploration of IT advantages: it is to explore advantages of IT aiming at solving difficulties existing in the course of traditional instruction and then systemically analyze these advantages to conclude rules.

(3). system designs: it should solve difficulties, which cannot be solved by traditional instructional means, by using IT advantages. It includes the resource structure, display form and application mode.

(4). curriculum resource making and integration: it is to select suitable developing technology and tool to make resource and make effective integration according to resource structure.

(5). resource application and feedback: it is to apply structural curriculum resource in the course of instruction, analyze the performance of these resources through dynamic instructional feedback, and then decide if it needs to repeat step1 to step5 or some step(s) of them.

(6). end.

It can be drawn from upper steps, ERDA needs a process, in which teachers, educational experts and IT experts explore and study together. ERDA needs practicing. It cannot be implemented by empirically simple reasoning. The nature and rule of ERDA can be discovered through instructional practice.

2. Curriculum Diagnosis Approach

Curriculum diagnosis is the base of ERDA. The quality of resource design and development is subject to the precision of curriculum diagnosis. Steps in curriculum diagnosis are as follows:

Step 1: classify curriculum content according to their commonness rule, take contents that have commonness rule as the same category. Some categories can be classified to sub-categories or multiple sub-categories.

Step 2: find suitable approaches of instructional mode design, instructional process arrangement and instructional strategy selection for each category.

Step 3: analyze difficulties existing in the course of general instructional activities and classify these difficulties into categories.

Among these steps, instructional mode and process design are the most important step. If they are not reasonable and scientific, the diagnosis results are unacceptable.

In the course of instructional mode and process design, some factors should be considered. The following are the key factors:

(1) what are instructional purposes? Where do students' learning motivations come from and how to inspire them? Students' learning motivations come from their needs of life instinct and curiosity. The learning purposes are to educate abilities, improve diathesis and master knowledge and information. Concretely, they include mastering not only basic knowledge but also corresponding knowledge system, logical thinking approach and essential information based on the frame of abilities. Thus, from the angle of psychology and education, learning contents that are suitable for the needs of students' life instinct and curiosity should be arranged for students.

(2) factors on affecting learning efficiency are multifarious, thereinto, age and information form are the most important factors among them. Different ages need different learning modes. People's means for obtaining information are mainly from five sense organs. Information forms affect directly the received efficiency. Thus, from the angle of psychology and education, the effective learning environment, reasonable instructional process and mode that is subject to the age character should be made for students.

(3) factors on relationship between knowledge education and innovation education should be considered. On knowledge education, its goal is to master knowledge and then change them into skills. Normally, Students educated in this way just can do the repeated work. However, society needs persons with innovation ability. Thus, the instructional result should be to make students think about more questions instead of no questions.

4 Information Environment Classification and Feature Analysis

The role that IT plays in the course of curriculum implementation is subject to the sufficient analysis on IT advantages and features. Analysis on IT advantages and features should be from multi-angles such as in which parts of instruction IT should be applied, what kinds of learning environment can be made by IT, etc.

The advantages and features of IT include dynamic interactivity, independent usability, from stillness to movement, from abstract to concrete, from micro-state to normal-state, virtuality, simulation, fast transmissibility, real-time, amplitude, etc. These features are good for inspiring students' learning interest, integrating excellent teachers' and experts' wisdom together and inheriting effective instructional modes and approaches.

In the course of IT and curriculum integration, we should do more research on IT method. IT can set up multifarious instructional environment. Instructional environment is divided into the following types: multimedia classroom, computer classroom, campus network and internet. It is necessary to study their features and application. For each instructional environment, we analyze it from the following aspects: instructional modes, instructional steps and needed condition, etc.

4.1 Multimedia Classroom

(1) instructional contents: cognitive learning on knowledge and question; explanation on question, method and process; etc.

(2) instructional modes: teaching class by using multimedia resource, interactive activities between teacher and students.

(3) learning steps: scene design, teaching, thinking and discussing, concluding、exploiting thinking skill, inter-subject learning, subject system , training, simulation test.

(4) needed conditions: instructional resource database, preparing for class platform and teaching class platform

4.2 Computer Classroom

(1) instructional contents: cognitive learning on knowledge and question; explanation on method, process, application and practice; ability cultivation; etc.

(2) instructional modes: teacher direct students self-learning or cooperative learning or research learning by network; network is not only display means but also support means for students' learning and testing.

(3) learning steps: communication among students or between teacher and students; real-time test and feedback; evaluation by students or teachers; scene design; teaching class; thinking and discussing; concluding; exploiting thinking skill; inter-subject learning; subject system; training; simulation test.

(4) needed conditions: network instructional resource database, instructional platform, self-learning platform, cooperative learning platform, instructional test and assessment platform, instructional management and feedback platform, research learning platform, learning resource website, etc.

4.3 Campus Network

(1) instructional contents: knowledge learning, review, stability and exploration; investigative learning ; stability, improvement, exploration, application and practice on approaches and process; research ability cultivation; improvement on students' integrated diathesis.

(2) instructional means： enlarging learning space and content by network, perfecting the existing subject system, exploiting field of vision.

(3) learning steps： communication among students or between teacher and students; feedback, stability, improvement and exploration; perfecting the existing subject system; inter-subject system, training system.

(4) needed conditions: network instructional resource database, self-learning platform, cooperative learning platform, instructional test and assessment and feedback platform, research learning platform, learning resource website, etc.

4.4 Internet

(1) instructional contents: knowledge learning, review, stability and exploration; investigative learning ;application and practice of approaches and process; ability cultivation; improvement on students' integrated diathesis.

(2) instructional means： enlarging learning space and content by network, perfecting the existing subject system, exploiting field of vision.

(3) learning steps： communication among students or between teacher and students, stability and improvement and exploration, inter-subject system, training system.

(4) needed conditions: learning resource website, communication support tools.

5 Resource Architecture

Definition 6: curriculum resource system means all kinds of resource types, which includes some based on curriculum criteria, some based on textbook version, some based on test resource, some based on public resource among curriculums and so on.

Definition 7: textbook version resource means instructional resources organized by some textbook versions. These resources mainly focus on resolving difficulties that general instructional means cannot solve during the course of instructional modes implementation. They are organized by curriculum unit, subject, chapter, passage, item and so on.

Definition 8: material granularity means knowledge and skill, process and method as well as the scale of sensibilities value implemented by these materials.

Definition 9: material share means materials that can be used in two or more than two courseware. There are two means for material share, one is courseware share for the same content, and another is courseware share for the different content.

Definition 10: curriculum criterion resources mean resources that are organized through basic knowledge system, encyclopedia information, special topics, test training questions, learning rules, teaching rules, etc. They don't depend on textbook's version. Some textbook's resources can be built based on them.

Definition 11: public resource means resource that reflect people's living environment and can be shared by subjects. It includes natural and humanistic resources. Typical public resources include information on nation, region, city, organization and people as well as some natural environment information on mountain, river, lake, ocean and so on.

Definition 12 test training resource means resource that can be used to test if students understand curriculum knowledge or not and to discover where problems exist. It includes curriculum forward model, curriculum backward model as well as related test questions and papers.

Figure 2 illustrates the structure of curriculum resources.

Curriculum resource can be divided into three hierarchies from figure 2, namely, basic hierarchy, category hierarchy, application hierarchy. Resources in basic hierarchy are the basic curriculum resource types; they have the maximal share degree, normally, they are 100 percent share. Resources in category hierarchy are integrated through specific questions; they have the higher share degree. Resources in application hierarchy are integrated for the concrete applications; they have the lower share degree.

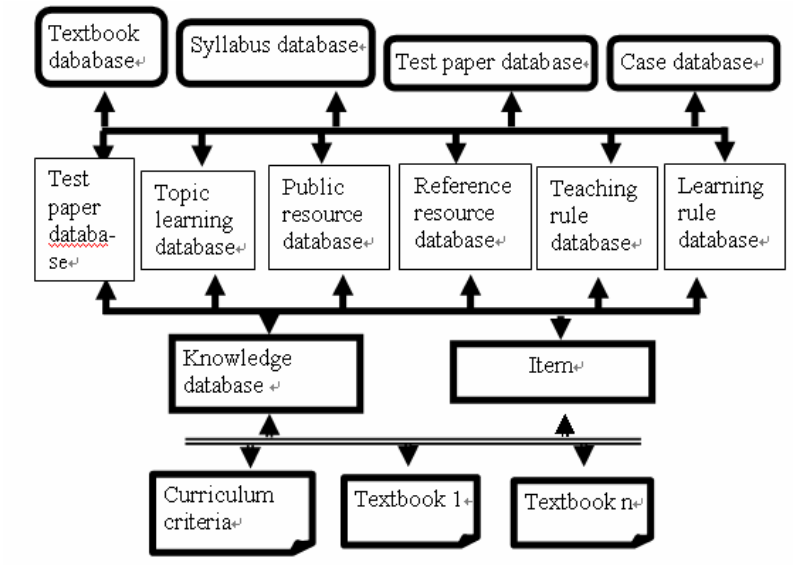


Fig. 2. Structure of curriculum resources

6 Curriculum Resource Evaluation

6.1 Resource Evaluation Factors

Resource evaluation should be considered from multiple factors. The following are key factors:

(1) the proportion of usable resource in each class;
(2) the proportion of share resource in each class;
(3) the granularity of share resource;
(4) the distributing uniformity of total resources;
(5) the reliability of resource design
(6) the validation degree of resource used in instruction;
(7) the freshness of resource update.
(8) the satisfaction degree between resource and curriculum instructional needs;
(9) the leading degree of resource.

The function of each factor is different, some of them are sufficient, some of them are necessary and some of them are sufficient and necessary. Each factor should be assigned to a weight in the course of evaluation.

6.2 Evaluation Criteria for Curriculum Integration

Evaluation criteria for curriculum integration include improvement degree of students' learning effect and operation performance of curriculum implementation. Concretely, they include the veracity of instructional goal, content, emphasis and

curriculum types; the scientificaalness of selecting instructional mode as well as arranging process, strategy and method for each curriculum type; the veracity for determining difficulties existing in the general instructional means; the perfectibility and adaptability of applying IT; the systematicness , effectiveness and feasibility of resource and software design; the operation performance of instructional resource, software and website and so on.

6.3 Evaluation Example for Network Curriculum

A. Factors

(1) the understanding of curriculum goal is correct or not r1
(2) the curriculum content determined by curriculum goal is reasonable or not r2
(3) the three-level directory is reasonable and logical or not r3
(4)the content of each item is balanced and enough for questions, system and multimedia materials or not p1,r4
(5) the learning direction and suggestion are sufficient or not p2,r5
(6) the reference material enable to meet all students' needs or not p3,r6
(7) exercise and thinking question are suitable or not p4,r7
(8) the interface is beautiful, the connection is right and the set up type is normative or not p5,r8

where，$0 \le r1$，r2，r3，r4，r5，r6，r7，$r8 \le 1$，and $r4+r5+r6+r7+r8=1$；
$0 \le p4$，p5，p6，p7，$p8 \le 100$。

B. Formula

After getting all information, compute the total score by using the following formula,

$$0 \le S \le 100 .$$
$$S=p1 \times p2 \times p3 \times (p4 \times r4+p5 \times r5+p6 \times r6+ p7 \times r7+ p8 \times r8) \quad (1)$$

C. Weight Assignment

R4=0.7, r5=0.1, r6=0.05, r7=0.1, r8=0.05

7 Platforms for Resource Management, Development and Application

Definition 13 instructional support platform software provides necessary support for instructional process. It includes resource management and share, preparing for class platform, instruction platform, management platform, local region communication platform and blog support platform.

Concretely, resource management and share platform includes resource management, retrieval, selection, delivery and share tools. Preparing for class platform includes edit, integration, and animation making tools. Instruction platform includes teaching, learning, test, training, assignment, answering question, communication, evaluation and

forum tools. Management platform includes daily management, electronic government affair and information publish tools. Local region communication platform includes information publish, information feedback and communication between teachers and students or between teachers and parents tools. Blog support platform includes support tools for teachers' communication, forum and blog. Figure 3, 4, 5 illustrate more details.

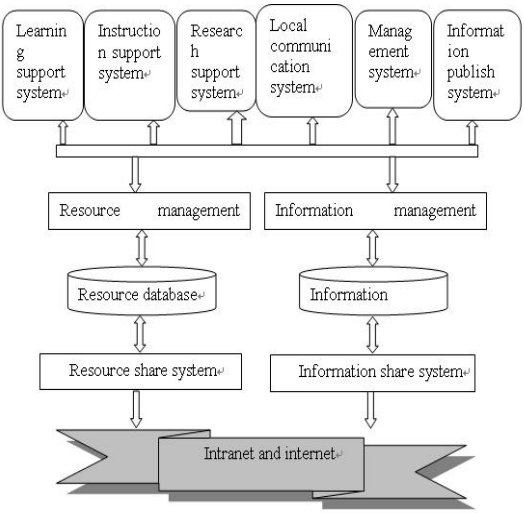


Fig. 3. Instructional support platform

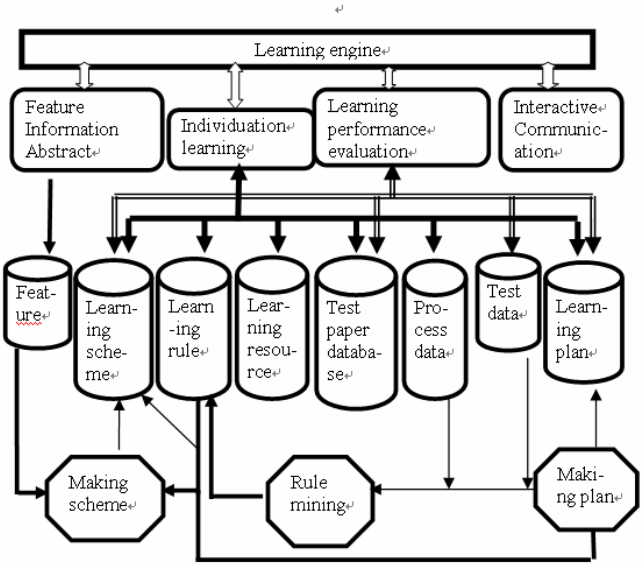


Fig. 4. Learning engine

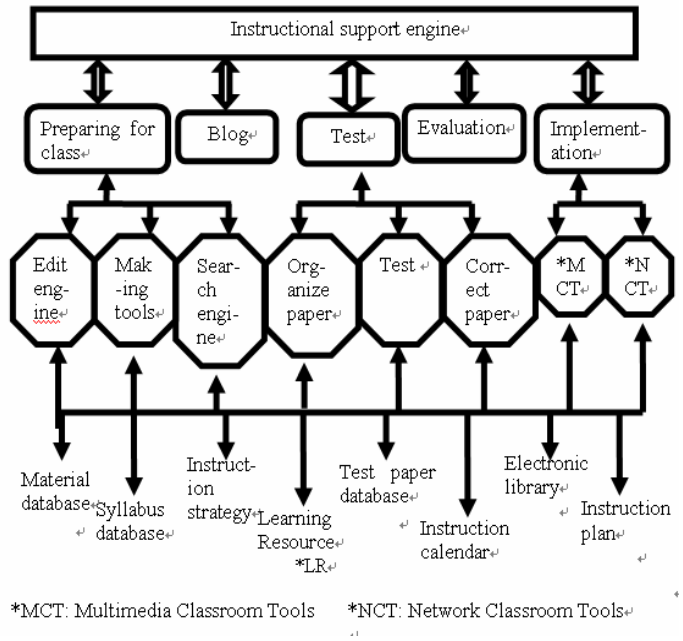


Fig. 5. Instructional support engine

Definition 14 software tool on educational resource making, which can help teachers to do educational resource by themselves even if they don't know how to program. Normally, there are two types: one is for making animation resource; another is for integrating some resources together.

8 Conclusions

In the course of ERDA, it should be to make clear the orientation; study curriculum types and difficulties existing in instructional process; arrange learning contents that are suitable for students; select learning process and environment that are helpful for students; cultivate logical thinking ability while applying visual technology; think about integration between subjects; establish teachers' continuing education system; improve application and management system; take integration as a system engineering.

Try to form a leading modes and methods system: modes and methods system on the integration of IT and different content type; on the establishment of resource and software structure; on the application of resource and software. ERDA needs educational experts, IT expert, officer and teachers work together to make great progresses.

References

1. Nan, G.: Several theoretical and practical problems on education informatization. J. E-Education Research 11, 3–6 (2002)
2. Sang, X., et al.: Strategy on district educational informatization. J. E-Education Research 3, 8–11 (2005)
3. Jin, Y., Zhong, S.: Analysis on modern long-distance instructional modes and support platform. J. Educational Technology Research 4 (2003)
4. He, K.: Nature of e-Learning–integration of IT and curriculum. J.E-Education Research 1, 3–6 (2002)
5. Gu, X., Zhu, Z.: Implementation mode on teachers professional development. J. China Educational Technology 3, 5–8 (2005)
6. Wu, J., Yu, S.: General problems and countermeasure for IT and curriculum integration. J. E-Educational Research 10, 70–72 (2006)
7. Tang, Q.: Analysis on errors of IT and curriculum integration. J. E-Education Research 10, 70–72 (2006)
8. Zhong, S.: Research on resource and software of IT and curriculum integration. J. E-Education Research 3, 53–56 (2005)
9. Zhong, S.: Model and method on IT and curriculum integration. J. China Educational Technology Research 9 (2003)

Research on Management of Resource Virtualization Based on Network

Gui-Lin Chen[1], Sheng-Hui Zhao[1], Li-Sheng Ma[1], and Ming-Yong Pang[1,2]

[1] Department of Computer Science & Technology, Chuzhou University, Anhui, China
[2] Department of Educational Technology, Nanjing Normal University, Jiangsu, China
{glchen,shzhao}@ah.edu.cn, panion@netease.com

Abstract. In order to satisfy the demands of individualized learning for various learners, it is necessary to reform the traditional resource management pattern to provide various learners an individualized learning platform and abundant resources. In this paper, the basic attributes of teaching resources are discussed firstly, and proposes to exert the resources using attributes by the virtualizing management. Thus, produces the definition of management of resource virtualization, and designs a layered management model of resources virtualization. Further, the article discusses the function and realization method of each layer, and proposes the concrete implementation suggestions. The developed partial practices developed showed that the virtual management could effectively promote resources sharing and expediently use for learners.

Key words: Resource Management, Resource Virtualization, Resource Sharing, Digitization.

1 Introduction

In practice of university education, it is becoming an increasing important topic to discuss how to deal with the problem about lack of resources. Especially, in recent years, those things, such as number of students in university increased remarkably and finance of education relatively decreased, bring a serous lack of resources in universities of China. The shortage of resources exists in many aspects, for example, equipments, locations and human resources, and so on. Obviously, increasing resources is a direct way to change the situation, but it is not the best selection for us since it can't economically solve the problems indeed. In fact, there are many resources, which are not used efficiently and sufficiently, in the universities. The main reason for this is that the resources mentioned before are exclusively occupied by different universities, which results the resources can not be shared with other universities. So it is an effective and natural way to solve the shortage of resources by reforming traditional pattern of resource management based on information technology, which can share resources and enhance efficiency of resource using.

The main task of university is to cultivate students and satisfy various demands of learners. The goal of instructing student, the pattern of education and

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 302–309, 2008.

the frame of organization are deeply effected changed by the coming information age. In university, Teaching environment rapidly improves and the limits between the different schools and subjects are being broken. The educational pattern based on the teacher's instruction is lagged with the times in industrialized era. Today, we need a new educational pattern to satisfy the individualized demands of students. At the same time, we also need to offer some environments and resources of learning with relevant patterns and corresponding management system to satisfy the demands of students.

Based on the situation mentioned above, we have to reform today's resource management pattern of universities to fully display all the teaching resources and to satisfy the learners' individualized demands. In this paper, we proposes a strategy to realize the resource management virtualization based on the information technology. The rest of paper is organized as follows. Firstly, we discuss the value and model of resource management virtualization and the related work in Section 2, then give a general framework of resource virtualization management in Section 3. Some suggestions about detailed implementation of the model is proposed in Section 4, followed by our conclusion of the framework and discussion about the significance of management virtualization pattern in Section 5.

2 Sense of Resources Management Virtualization

In past decades, researches including virtualizing campus , virtualization enterprise and so on are paid more and more attention by researchers. As we know, the virtualization technology is created following the development of computer technology and Internet, and related concepts about it is also originated from the computer technology. According to the Wikipedia's definition,"virtualization is expressed the process of computer resources logical group (or subset), visit them like this way which obtains benefits from the primitive configure. This kind of new resources virtualization view isn't limited by realization, location or the first floor resources physical configure."[4]. Another kind of more direct definition is,"virtualization is logic indicated of resources, it isn't restrained by the physical configure limited." [4]. According to the above definitions, the concept of virtualization contains the following meanings:

First, the virtualization is a kind of abstraction about the actual physical resources and data resources. The virtualization resources is a logic expression of actual resources and it does not change the physical properties of the original resources. Second, the goal of virtualization is to reduce the coupling degree between users and the resources and to provide users's task an implementing way, which does not rely on the specific resources. Thus, the users can obtain the resources according to their demands, but does not care about the concrete position of resources and is not limited by actual property of resources. Third, the users may gain more benefits from the virtualization than traditional way. In other words, virtualization is did not consider the physical property of the resources, merely considered how to use effectively whcih cause the user to obtain

benefits. Considered from the user, virtualization is only considered uses, did not consider possesses.

By analyzing the kinds of teaching resources, three kinds of attribution should be found. The first, attachment. That is to say, which department the resources belong to. Certainly, in view of overall aspects, the school has all ownerships. But according to current circumstance, it usually still belongs to a certain section in the school. The second, management . The resources needs to be managed and maintenance, this attribute usually has relation to attachment, in anther words, it is managed and maintenance by whom it belongs to. The third, usage. It is the attribute that serve the teaching and play an important role. This is the most important part, it is also a part that customer to compete with each other fiercely. in current circumstance, three kinds of above-mentioned attributes of teaching resources are generally and all managed by a department, it is also the traditional management style that we usually say. the share of resources is limited by this kind of mode, because proprietor of resources usually provide a service toward other customers after satisfying own demands, but if the usage application of other customers conflict with the department benefits, the proprietor will carry on restriction to the resources share.In this kind of mode, all resources have to unify together.It can't provide the learner individuation study space, and it can't satisfy the learner's study need.

Management of resources management is putting forward for resolving the above mentioned problems, the purpose is to maximize the usage attribute, and provide individuation study environment for the learner.

So we can give such definition: management of resources Virtualization of the university and college base on the foundation of network, on the premise of keeping management and attachment attribute unchanged, usage attribute may be separated from other components ,constitute a logic resources layer(or a virtual resources layer), and build up a specialized resources Virtualization scheduling center to schedule resources, according to the certain management strategy. Any user can apply for logic resources to handle his mission.

3 The Resources Virtualization Management Framework

According the status of the resources management and technology, the paper designed a six layers model for resources virtualization management as Fig.1. The middle four layers is transparent for users and physical resources. For users, the integration and scheduling is completely transparent for users. For resources, what they see are users, and they do not care the dispatch or who to dispatch.

3.1 Physical Resource Layer

The physical resources are the real teaching resources, which can be laboratory, classroom and other physical resources and also they can be electronic literature, network classroom and other numeric learning resources or human resources.

The relegation of the physical and the manage property is invariable. The manage department of the school and the relegation department of the resources,

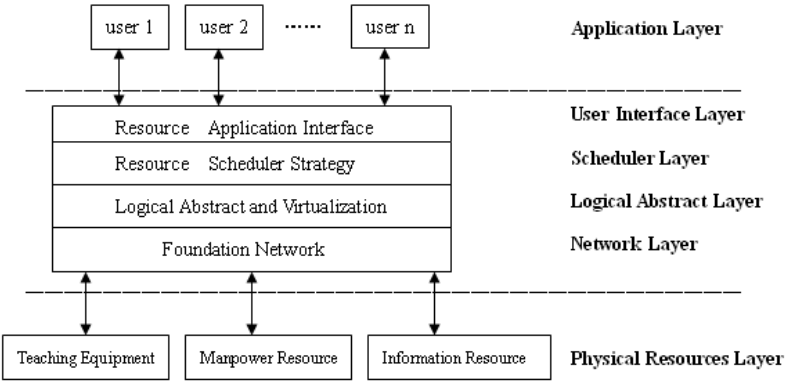


Fig. 1. The Model of Resources Virtualization

according the strategy of the school and the fact instances, can decide the resources whether can be shared, or the range of sharing.

3.2 Network Layer

It is the network infrastructure of the campus. A stable and reliable campus computer network is the precondition of the realization, because the virtual resources management is based on the information technology.

3.3 Logical Abstract Layer

It is the core layer, which is responsible for supervising the state of the physical resources, and collects the information of shared resources. It abstracts and integrates the shared resources and virtualized them. According the different character of the physical resources, the virtualization can be classed into three types. The first type is information resources, eg, computation , storage and electronic data. They can be integrated and virtualized by grid and web services technology[4,5]. The second is classroom, laboratory and human resources. The main problem of these resource is the scheduling, and it is simple to implement. The third is the virtualization of the great experiment instruments, which can be implemented by network and other information technology. The logical abstract of the resources can be multiplex. A single resource can be expressed into more than one logical resources, or some resources can be expressed by one logical resource. The detail virtualization strategy is not discussed here.

3.4 Scheduler Layer

It is a digital resources scheduling center, which is very important. It is responsible for supervising the resource state of virtual resources. It receives the user apply from upper layer, and dispatches the logical resource for user according the established schedule strategy.

The resource scheduler strategy can be planned by the management layer and the owners, and integrated into resource scheduling layer by the technician. The resource is automatically dispatched according the established scheduling strategy, which is fair for the users. The results is submitted to user by user interface. If the shared resources are the really monopolistic physical resources, it must inform the resources management department. While, it is more complex to implement for integrated resources, while can be settled by technic layer. The users and the management layer do not to care.

3.5 User Interface Layer

Providing the user interface, a portal. It includes some shared information resources, user login interface, and so on. It is the only interface for user applying for resources. Actually, It is users who apply for resource using from portal.

3.6 Application Layer

The main object of this layer is users.

4 Implement Suggestion of Virtualized Management

Realization of resource sharing has become the consensus of the majority higher education workers, but in the specific practice, there are still many factors affecting share. Aim to the management model of the virtual resources present in this paper, we put forward the following proposals for its implementation.

4.1 Change Concept

The most important is still the concept and it is need to promote the concept changed by various ways such as propagation and education. Judging from universities' reality, we know that majority branches are still habit monopolizing various resource in self, especially the teaching unit. Except the factor of management, concept is a very important factor. We deal with several viewpoints here.

Firstly, school leads' concept. Come to say resource's the sort , lead to be to be ready to come true tier. But in putting process into practice concretely, still existence some affect the factor leading a wish's. For instance, the management boundary between different administers would be mixed up because of share. In other words, it is hard to find junior organization. Obviously, it is a question of habit. If we do canonical management, these questions will not appear on the way of our work.

Secondly, functional branches' concept. Functional branches always think that the management is its patent. Secondly, functional branches' concept. Functional branches always think that the management is its patent., the virtualization management will affect the management way of the functional branches without a doubt, even will be able to form the feeling which one kind of management

jurisdiction will reduce. If this level cannot change their idea, it will seriously affect the virtualization management implementation. Thirdly, branch of resources ownership and user's idea. In the university, user and ownership branch's role are mutually transform in some time. In the tradition, everyone is accustomed to use monopolistically.

4.2 Gradual Implementation of Systems Layout

The core content of resources virtualization management contains the following three point. First, resources sharing; Second, separates the use rights and the property rights, then make the use rights with uniform management and distribute. Third, satisfies the user's individuality demand. In the implementation process, according to the resources condition of the school, we can frame the system implementation plan, determine the resources that bring into the virtualization management, design the resources distribute strategy. Based on this, we can gradually implement according to plans. Moreover, we need fully investigate and proof the existing resources and its running condition in the university, then carry out the difference treatment according to the different type of the teaching resources. The different type resources should have the different virtualization strategy and management distribute strategy.

4.3 Pay Equal Attention to Technology and Management

In the implement process of virtualization management, the virtualization management has two point of support, involving how to make the virtualization resources and how to manage, which is indispensable,. Therefore, we not only highly attach importance to the technical method, because the virtualization resources can not establish without the technology foundation, but also highly attach importance to the management measure, because even if the virtualization system established, it also could not normally run without the management safeguard.

Considered from the technology, the special technical team must be established based on the school information organization, and the technology scheme of resources virtualization must be constituted. Considered from the management, the management must be criterion, the responsibility of different management level must be explicit.

The traditional multi-level management is needed to transform to the flat management ,and the management level is reduced, which involves function adjustment of the original management level and the management branches, even involves the benefit redistributes, therefore also which must be implemented on the premise that idea is changed.

4.4 Complete the Specialized Division of Labor

On the other hand, in the concrete virtualization management implementation process, the various role which involves the user, the resources superintendent,

the people of resources integration, the resources distributer and so on need to have the specialized division of labor and clear about its jurisdiction responsibility.

Finally, the information is the foundation of virtualization. The infrastructure.of school's information must be consummate for implementing virtualization management.

5 Conclusion

The pattern of the resources virtualization management which designed in this paper based on the existing resources can enhance the efficiency of using resources and sharing resource, through the information and standardized management, also can provide the individuality study space and the environment for the learners. From the experiment which already developed see, which had the following several mean.

5.1 Enhance the Efficiency of Using Resources

The efficiency of using resources can be enhanced through the virtualization management. Take the efficiency of using computers in Chuzhou university 's laboratory as an example, according to the statistic, before the management pattern reformed, each machine daily averagely uses 3.8 hours, after the virtualization management implemened, the daily average use increase to 6.9 hours, the efficiency enhanced about 80.

5.2 Promote of Individualization Study

The learners can apply the learning resources expediently with their requirement based on virtualization management environment. From the reality of our school, the network teaching system works for 18 months, during this time, there is more than 1100 free users to open classes which are not included in teaching study, and the number increased to 1500 after opening the virtual experiments three months later. It illustrates that virtualization management promoted the initiative study of the students. Of course, there may be cases that the students login only for their curiosity.

While it is adopt the virtualization management, the existing management pattern will inevitably change to corresponding flat management pattern and it will become more efficiency to school teaching.

Moreover, the virtualization management of resources is advantageous to promote transformation of college management concept and form the opening environment of universities. Fondly speaking, it is used of students in culture.

Obviously, the virtualization management pattern of resources proposed in this paper is an experimental model. And it is need to do more research about management, technology and concept if the model uses in real world. In future, we will do further research with the practical situation of our school to the technology method, the adjustment of management pattern of virtualization resources and make it playing an important role in teaching practice.

Acknowledgement

The work presented in this paper is supported by the Natural Foundation of Anhui Provincial Education Department (No.2006KJ041B and No.KJ2007B073), the Teaching Research Key Project of Anhui Province of China and the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (No. 2007jyxm115 and No. 07KJD460108).

References

1. Trow, M.: Lifelong Learning through the new information Technologies [J]. Higher Education Policy 12(2), 201–217 (1999)
2. Trow, M.: Some Consequences of the New Information and Communication Technologies for Higher Education[EB/OL], http://cshe.berkeley.edu/publications/docs/
3. Egol, M.: The future of higher education[J]. Educause review 41(4), 72–73 (2006)
4. Maldonado, M.F.: Virtualization in a nutshell: A pattern point of view[EB/OL] (June 2006), http://www.ibm.com/developerworks/library/gr-virt/
5. Garbacki, P., Naik, V.K.: Efficient Resource Virtualization and Sharing Strategies for Heterogeneous Grid Environments. In: 10th IFIP/IEEE Symposium on Integrated Management (IM 2007), Munich, Germany, May 2007, pp. 40–49 (2007)

The F-R Model of Teaching in Chinese Universities

Hui Zhao,Yanbo Huang, and Jing Zhang

Modern Education Technology Center, Central South University 410083, Changsha, China
huizh@csu.edu.cn

Abstract. In view of the teaching characteristics of the full-time universities in China, the paper probes into the application of web-based teaching platform and the new model of teaching--the F-R model of teaching that integrates the rigid teaching management with the flexible curricula teaching to make full use of the teaching resources and enhance the teaching efficiency. The new model of teaching not only assures the integrity of the students' knowledge system but also emphasizes the individual difference. It provides a way to cultivate students' personality and develop the students' ability to analyze and solve the actual problems that are necessary for the full-development and innovative person in modern society.

Keyword: Higher School, Instruction Design, Flexible System.

1 The Current Situation in Chinese Universities

In China, the student management in most of the full-time universities is in fact according to the credit system of school year. Under this system the students can only choose courses in the range that is designed rigidly by the Education and Administration Department. This gives the students very little freedom to choose courses. With the enlarging of the recruitment, the extending and incorporating of the universities, many universities present multi-campus pattern that teachers and students are in different campus. Though the amount of the students is increasing fast, the amount of the teachers isn't increasing correspondingly. In order to accomplish the scheduled teaching task, the teachers often have to take on more than one course simultaneously. Or several classes attend the same class together, namely so-called "big class" in Chinese. In this case, as the large amount of the students and the increase of the work, the intercommunion between the teachers and students decreases. Correspondingly, the teachers know little about the study condition of the students and the teaching content is separated from the demand of the student cognition.

In the classroom instruction, most of the teachers follow the traditional model of teaching and adopt the one way process that the teachers teach and the students listen. The teaching reform rests on the surface of the modern education technology that is the teaching equipments such as projector、projector slice and audio-video documents are used as auxiliary tools to help teachers lecture for the multi-class. In this case the modern education technology is utilized merely to help display the teaching content and lessen the teacher's repetitive labor. Under such model of teaching, because of

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 310–315, 2008.

negligence of the students' cognition process, the students are in passive situation and can't carry on deep thought. So their thinking mode and innovative abilities are fettered.

The information technology instruction that is grafted on the traditional model of teaching consolidates the traditional model of teaching and doesn't make the real modern education technology instruction come true. Thus in order to promote the fusion between the information technology and the instruction in university and improve the quality of instruction in university, the paper tries to make some explore on the model of teaching in Chinese universities and puts forward the F-R model of teaching for the first time. The new model synthesizes the rigid instruction based on class in universities and the flexible instruction based on education of individual information accomplishment. And the new model makes the realization of instruction according to person that based on learning task come true with the help of information technology and web-based teaching platforms. In this case, the students can study independently and construct their knowledge meaningfully under the teacher's instruction. So it can help to develop the students' innovated ability.

2 Analysis on the Model of Teaching

Education is an activity that exerts an influence on the morality, the intelligence, and the physical of educatee based on certain request. In essence it is a kind of activity to cultivate people. In order to adapt the social demand and the development of science and technology level, many models of teaching have been discussed. In conclusion there are two kinds. One is "the flexible model of teaching" that emphasizes the individuality of the students; the other is "the rigid model of teaching" that emphasizes the improvement of the teaching efficiency.

2.1 The Flexible Model of Teaching

About more than 1000 years ago there was "Si Shu" that the teacher exerted individuality education on the students in China. That the artisan guiding the prentice was another pedagogical activity based on the study skill. Both the "Si Shu" teaching and the apprenticeship teaching may belong to flexible model of teaching. The characteristic of flexible teaching system is that under teacher's instruction students can make asynchronous study come true. Or in other words, the content and the progress of the students may be different. The teacher manages the whole process. According to the diversity and the desire of students, the teacher devises different plan of instruction for different students and adjusts the teaching goal. The flexible model of teaching can reach multidimensional cultivation so the students can devote themselves to the development of society more quickly.

The obvious characteristic of the flexible model of teaching is to implement individual instruction to the students. It is good for the teachers to discover students individuality, instruct the students to study and develop their personal thinking. Under this model of teaching, the teacher plays an important role in the improvement of the teaching equality. The teacher manages each student's entire studying process; therefore the amount of students instructed by each teacher is limited. So the flexible model of teaching is not adapted to popular education.

2.2 The Rigid Model of Teaching

After entering the 18th century, the assembly line started to be applied to the mechanical industry to make large scale production. Correspondingly, the same characteristics come up in the pedagogical activity. That is to divide the teaching content into certain units: required curriculum, technical required curriculum, professional required curriculum and professional curriculum. The process of instruction was similar to the assembly line. After passing the “assembly line", the students turned into professionals. This kind of model enhanced the teaching efficiency enormously, so a large amount of professionals may be “produced” in a short time.

The rigid model of teaching emphasizes the collective design of instruction. Teaching plan is made according to most of the students in class and the teaching goal is explicit. So it is easier for teaching management. The rigid model of teaching exerts level- teaching according to the structure of knowledge so it is easier to shape the thinking process of the students. While the disadvantage is that it ignores individual difference so that one part of the students can’t exert their potential adequately and another part of the students deal passively and can’t acquire knowledge in true sense.

2.3 The F-R (Flexible—Rigid) Model of Teaching

With the application of multi-media computer and Internet in education, it builds a good teaching reform platform for improving the teaching efficiency and teaching quality. In order to adjust the large-scale and the cross-school-area teaching in full-time universities in China and apply modern education technology to new teaching system, it is imperative to study new model of teaching to cultivate innovative person. The primary content of the F-R model of teaching is to introduce the flexible teaching idea emphasizing individuality into the rigid model of teaching management based on class in the full-time universities. With the assistance of the modern education technology to the Education and Administration Department, the new model manages students centered on course and provides independent learning environment for students. During the curriculum teaching, with the help of modern education technology, the teachers can teach students based on individuality of the students, help students improve their information accomplishment and explore new knowledge and make asynchronous study come true.

In the F-R model of teaching, modern education technology is utilized during the entire process in the teaching management: during the primary stage to learn elementary knowledge, modern education technology presents the study task to students so that students are clear about their study direction. With the deepening of their study, the desire of their study task is enhanced gradually and at last the goal to grasp the knowledge of their major is achieved. During every stage of their study, with the aid of the modern education technology and the instruction of the teachers, students can not only select course to attend, study independently, broaden aspect of knowledge but also take part in scientific research and acquaint themselves with the latest development of their major.

With the aid of modern education technology, on one hand, the F-R model of teaching strengthens the interaction among the Education and Administration Department, the teachers and students; on the other hand, it helps students to manage

their study resources and is convenient for teachers to instruct students to study independently, to explore new domain and to develop students' innovative ability.

3 Design of the F-R Model of Teaching

3.1 The Goal of Design

Presently the teaching system emphasizes the influence to students from the outside and the one-way transmission of knowledge from teachers to students in the full-time universities in China. The teachers make the teaching plan according to the traditional model of teaching, and teach the class according to the uniform teaching plan with the isolated teaching methods such as multimedia. During instruction, the teaching progress is made according to the students in the medium of the class without considering the difference between students. So it is difficult to cultivate innovative people that are demanded by society.

According to the characteristics that the teacher teaches based on subject and the students study based on grade and major in the full-time universities in China, the F-R Model of teaching pays more attention to the students' meaningful construction under the instruction of the constructivism theory. Also it establishes cognitive environment based on the learning task and multi-level network knowledge system within school for students to develop the learning ability under the instruction of the teacher.

3.2 The Management of Educational Administration

The management of educational administration is a model of teaching management that integrates the rigid management with flexible study of the students based on class. Maintaining the traditional teaching class system, the teaching content adopts the gradient model: introduction to the major-elementary knowledge-major knowledge-application knowledge. During study, the students establish the knowledge network based on subject so they can study and browse related knowledge with the help of the guiding system. The freshmen begin to study their major conspectus course when enrolling school and set up their major background so that they can cross the threshold and make sure the learning task earlier. The instruction is leveled as required curriculum, technical required curriculum, professional required curriculum and professional curriculum based on subject. The outcome of the student's study is evaluated synthetically with the quality of the thesis or the graduation project.

The flexible model of teaching is adopted to teach the knowledge of every level to students. According to the teaching goal of every subject, the teachers make teaching plan including the theory knowledge and the practice knowledge by stages and construct major knowledge network to help students study. Students can learn not only in accordance with the path designed by teachers but study independently integrated with their own desire. In addition, it permits students to choose course cross grade according to their own ability. For example, the related courses can be integrated as one course to teach. The detailed knowledge referred to can be decompounded to teach in major course so as to instruct students to practice.

3.3 Teaching Process

The teaching environment of the F-R model of teaching is the multimedia classroom and the web-based teaching platform. In the rigid teaching, the teaching carries on in the multimedia classroom with all of the students as the teaching object and the direct interlocution between the teachers and the students as the main alternating way. It concentrates on assigning study task, exploring the study methods to enlighten the students' creative idea. While the flexible teaching processes on the web-based platform. It constructs knowledge network with a great amount of digital resources to provide the hotspot knowledge for students. Thus, it not only promotes students to study innovatively but also enhances the teaching efficiency greatly.

The F-R model of teaching can also be applied to evaluate and assess the students. The rigid measure of the written examination can be used to appraise the students' ability to grasp the elementary knowledge. The flexible measure such as the discussion problem in the open-book examination or the writing of paper can be used to assess the creative ability of the students. With the integration of the two ways, the students' level of grasping their major knowledge and their ability to study independently can be evaluated and assessed synthetically.

The F-R model of teaching integrates the rigid teaching management with the flexible teaching method. Based on the major and the grade, the teaching management is divided into the establishment of the teaching object and the teaching content, the real-time teaching, the practice online after class and so on. In each teaching stage, the flexible teaching can carry on with the assistance of the web-based teaching platform.

4 The Application of the F-R Model of Teaching in Instruction

In the teaching process in class, the teachers teach the key points, the difficult points and the doubtful points under the instruction of the teaching goal and considering the students' actual status to guarantee the integrality of the knowledge system and concentrate students on the main points. Also the teachers design the knowledge network including the history of the course, the characteristics of the course, and the applied domain in society and so on to help students build a learning framework and make them have a sensible cognition of the course. So that students can have a definite object in view in their later study. Dividing the course into units and put them in the master nodes of the curriculum knowledge network, with the aid of guiding system, it can help students study asynchronously and nonlinearly.

The constructivism theory proposes that the study is related to "scene". Through the medium function of "scene", it can stimulate students' association effectively and arouse the related knowledge, the experience and the idea of the student's original cognition structure. It can help the students to assimilate or accommodate the new knowledge so that the gap between the knowledge and the solution can be reduced. During the teaching process, the teachers should try hard to create more genuine "scene" to lead students to study with the genuine task so that student can achieve sensible construction. With the "scene" driven study, it can not only encourage students to explore and study the problem from more than one angle but also is good to the transplant of knowledge.

On the foundation of the independent study, the students can carry on group discussion under the instruction of the teachers. Through this process, students can

discuss and discriminate mutually based on the “scene”. Thus students can achieve their sensible construction and develop their high-level cognitive ability.

Each group introduces the difficulty and the attainment in the exploring process briefly to the whole class. It can both facilitate other students to study, oppugn and evaluate and activate the atmosphere in class. That the teacher concludes in time can help students systematize the scattered knowledge they obtain effectively. And under the hint and the introduction of the teachers, students should try to conclude and summary by themselves.

4.1 The Online Practice and Q/A Stage After Class

That the assignment, the submission and the rectification of the homework are an important alternant means to communicate the teaching of the teacher and the study of the students. Through the practice, the students may consolidate and enhance the comprehension and grasp of the learning content. During the practice and the review after class, students may meet difficulties and need help and support at any moment. The web-based teaching platform provides convenient accesses to solve these problems and makes study more open without the limit of the space and time.

The Q/A system is divided into the synchronized Q/A and the asynchronous Q/A: The teacher may ask the students synchronously and also may design the intelligent teaching system to answer the common questions. The teacher can further obtain the study state of the students while answering. It is good to the flexible teaching, too. At the end of the study, the quality that students complete the task can be evaluated by “the work” —the thesis or the project of graduation

5 Conclusion

Under the F-R model of teaching, the advanced instructional and educational thought is taken to the class to instruct the whole instructional and educational process and cultivate innovative person. In the actual teaching process, the teachers should appraise every student’s study condition and adjust some links to meet the actual needs. Because the F-R model of teaching integrates the classroom bearing the emotion exchange with the free and open learning environment based on information technology, it both exerts the teachers' domain role and emphasizes the students' main body status. So it forms the harmonious atmosphere in the teaching process and enhances the study effect. Thus it guarantees the cultivation of the students' emotion, attitude and values and shapes the students' perfect personality

References

[1] Hui, Z., Xiangjun, C.: The interactive way and application of web-based instruction [J]. Educational Technology (2004)
[2] Kekang, H.: Constructivism —the theoretical foundation to innovate traditional teaching [J], E-education Research (1997)
[3] Shizhen, D.: Modern Education Information Technology Basic [M]. Central China Normal University Press, Wuhan (2000)

An Approach to a Visual Semantic Query for Document Retrieval

Paul Villavicencio and Toyohide Watanabe

Department of Systems and Social Informatics,
Graduate School of Information Science, Nagoya University
{paul,watanabe}@watanabe.ss.is.nagoya-u.ac.jp

Abstract. This paper presents an approach to design an interface for document retrieval, based on techniques from the semantic web combined with interactive graphical features. The purpose of this study is to enhance the user's knowledge while he/she browses the information through a graphical interface. In this paper, two aspects are considered: Fist, interactive features such as object movability, animation, etc. are discussed. Second, a method for visually integrating the search queries and the query outputs is addressed in order to retrieve documents. The visual features and the querying method are combined taking account the semantical relations among extracted information from the documents. This combination is evaluated as a means to determine the most suitable location for the results inside the interface.

Keywords: Visual Interface, Semantic Web, Visual Query.

1 Introduction

In a learning environment, it is important to allow the discovery of new knowledge. Based on this aspect, learning support can be added to different activities such as the information search. User support systems for searching interesting information have been studied. In these systems, user interaction monitoring is conducted [1]. However, it may be complicated to track the user's intentions in applications which contain common graphical components, such as drop down menus, list boxes, etc.. Graphically rich environments (for example those that allow free object movement and arrangement) may provide a more effective way to keep track of user's actions.

Another aspect of learning environments is the source of knowledge, which usually consists of databases, and large document repositories. Knowledge is obtained from these repositories by document retrieval methods. These methods have been developed from linguistic analysis, statistical methods, artificial intelligence, etc., in addition to techniques from the semantic web [2], such as the use of ontologies for document indexing [3,4]. Although the techniques of document retrieval become more precise in order to take full advantage of their potential, the application of user interface that can abstract the querying process is needed. Visual queries can be used in computer learning environments. Users,

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 316–323, 2008.

in computer learning systems, may not have the technical skills to execute complex command queries. Visual queries have been applied to ontologies mostly as search and construction tools [5].

In this paper, we propose an approach to design an interface for document retrieval based on their semantical information. Our approach consists of three parts. First, we discuss the use of ontological information in the indexing of documents. Second, we describe a method for translating the relative positions of graphical objects into semantic query statements and a means of integrating query objects with the query results in a common graphical area. The main purpose of this study is to develop a prototype system which enhances the user's knowledge while he/she searches information. Using graphical features, such as free object movements, queries can be executed dynamically, providing information to the user as he/she interacts with the visual elements.

2 Related Works

Document retrieval methods using semantic web technologies such as ontologies have been popular in recent years because the procedures for semantic web involve textual analysis. These techniques are also used for information extraction in document retrieval systems [6,7,8]. The system KEA [9] uses a method of recovering documents through the use of ontological information. Initially it applies several textual procedures such as stemming, common word removal, and document indexing. The indexes are created with the use of an ontology. KEA was used in this research because of the good extracting results in comparison with system-extracted and expert-extracted keyphrases [7].

Due to their hierarchical structure, it may be difficult to visualize the ontologies. Graphical tools have been developed for a better understanding of ontologies [10,11]. These studies use several graphical approaches in order to simplify the amount of information to be displayed

Graphical interfaces for document retrieval have been specifically researched [12,13]. Because retrieving documents can create extensive lists of results, specially in large repositories, it is necessary to apply visual aids to recover such information. Although researches exist in the fields of document retrieval, visual interfaces and ontology visualization, a combination of these technologies has not been approached extensively.

3 Framework

Figure 1 presents the framework of our proposed system. To realize the described approach, the following functions are needed: First, a function for extracting the most relevant keyphrases from documents, using information from an ontology; second, a function for executing queries over the ontology and the extracted keyphrases; third, a function to transform the queries and results into visual queries; and finally, a function to provide an interactive interface.

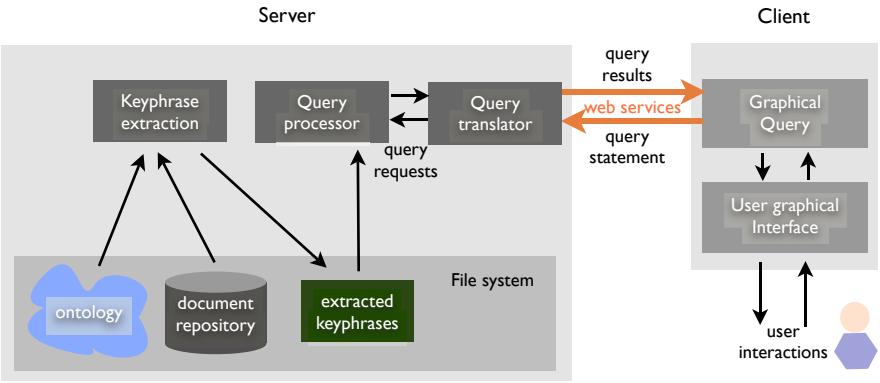


Fig. 1. The components of the server and client in layered framework

The framework is divided in two tiers, a web application server, and a desktop client. The web application contains the function of keyphrase extraction and query functions. The desktop client contains the interface functions. The communication between the server and the client is done via web services. Although the document repository and the ontology reside in the file system of the server, it is not restricted only for this server. File systems in other machines may be used as well, in order to create a more flexible structure.

3.1 The Keyphrase Extraction

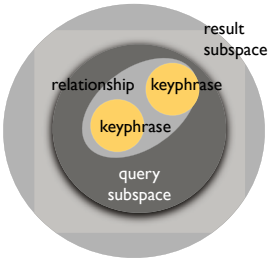
The digital documents consist of texts stored in the file system. These are first processed by the Keyphrase Extraction Algorithm (KEA)[3]. This extraction method requires an ontology. The textual documents are processed using lexical methods, and then analyzed with machine learning mechanisms, in order to select the most relevant keyphrases. This method is semi-automatic due to the fact that the machine learning requires a set of documents with manually selected keyphrases in order to build a model. Later, this model is used to process other documents.

3.2 Query Execution

Our framework utilizes CORESE (COnceptual REsource Search Engine) [14] as the query platform. CORESE is an RDF engine based on Conceptual Graphs. SPARQL is used as a basis for the query language and it also provides an inference rule engine. Therefore, it is possible to combine ontologies and data in RDF format within queries. CORESE has other functionalities that are not present in SPARQL, such as the approximated search (which finds best matches according to the RDF schema, path patterns, aggregation, etc.) This is accomplished by using the command `MORE`. Our system takes advantage of these added functionalities.

3.3 Graphical Representation and Query Transformation Function

In the graphical representation of a query, small circles stand for keyphrases and ellipses for relationships. The combination of these items contained inside a larger circle represent a query structure graphically. After a query execution, results are also illustrated as circles located in the surroundings of the query circle. Figure 2(a) shows the graphical representations and Figure 2(b) is a SPARQL query statement.



a) Graphical query

```
prefix skos: <http://www.w3.org/2004/02/skos/core#>
SELECT ?uri ?label WHERE
{
    ?uri skos:prefLabel ?label
    FILTER (?label ~ 'text')
}
```

b) SPARQL query statement

Fig. 2. (a) The graphical representation of a query statement. (b) A SPARQL query statement.

In the transformation process the SPARQL query statements are constructed. The graphical objects become query elements in which query subspaces shown in Figure Figure 2(a) are the main statement structure, and each keyphrase becomes part of the FILTER input in the SPARQL query statement shown in Figure 2(b). The arrangement of the results is graphed with the use of force-directed algorithms in connection with the elements that formed the query. This allows graphical links between the queries and the results. Relationships are defined graphically depending on the distance between keyphrases. Since the user can arrange the keyphrases in a desired way, the location of each keyphrase is calculated. If along the surroundings of a selected keyphrase exists another keyphrase, an ellipse with a center on both keyphrases is drawn.

3.4 Interactive Interface

The interface consists of elements that can be freely positioned inside the user interface layout. The locations of the elements provide important information to the system, since the query elements depend on their locations with respect to others. At each user interaction, such as element focus, element drag, etc. the system analyzes the manipulated elements in comparison with others. If the new arrangement changes elements in query, a query statement is formed and the information is sent as a query statement, results are immediately obtained upon the SPARQL query execution. Since results constitute of keyphrases, the

size is small enough to avoid communication delays. Therefore the process is of dynamic queries. The use of animation in the interface allows the display of the results as a dynamic process.

3.5 Data Structure

Between each process several data structures are constructed. In the process of keyphrase extraction, an ontology is used formated with the SKOS (Simple Knowledge Organization System) XML format. The SKOS format is a standard defined by W3C for representing ontologies. Also used in this process are the documents, which are presented as XML texts. After the extraction process has been completed, the result is a list of keyphrases and the corresponding documents, which is also in an XML format. This format enables CORESE [14] to process the queries. The results and the information from the query is communicated from the server to the client and vice-versa; and this is done through web services in which the data is in XML format.

4 System Characteristics

4.1 Visual Interface

The visualization of documents indicates the relationship among keyphrases extracted from the ontology. In our prototype system a query transformer and the interface elements are connected, so as to enable the user interface interactivity. Events occurred in our interface, such as changing the position between elements, are sent to the query transformer, which translates the event request into a query statement.

A query can be visually formulated by moving keyphrase objects inside a query space. The user can move these objects freely; therefore, they can be placed in

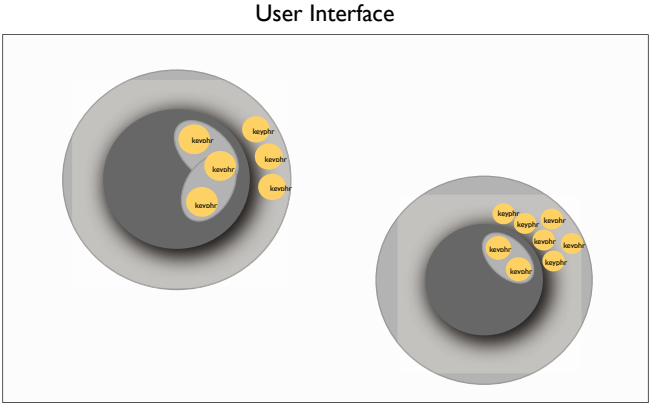


Fig. 3. A representation of the user interface

any location. The results from a query are returned after analyzing the locations of objects in the interface. Figure 3 shows the user interface with two queries. In this case a user has created two queries with related keyphrases. Results are displayed by the system in locations near to the keyphrases inside a query space. Each keyphrase object contains the location and the term of keyphrase. This information is processed by the query transformer in the following way: first, the positions of all the keyphrase objects and the query spaces are collected and sent to the query transformer. Second, with this information, each query space is transformed into a statement, using the keyphrases as search elements. Third, the query is executed over the ontology-document map stored in the keyphrase extraction phase. The query result consists of semantically similar keyphrases to moved keyphrase. Forth, the list of results is sent back to the query transformer, which translates the results again into graphic objects for the interface. The results are placed according to the position of the queried keyphrase and in relation with the query space. The process of transforming the visual keyphrase elements depends on their graphical positions. The query takes into account the relative distance between each keyphrase, inside a query space. The query transformer compares the distances between each keyphrase, if the distance is less than a defined value a relationship is added to the query statement.

5 Prototyping System

The visual interactive interface consists of two spaces: the query space and the result space, with a circular shape both will determine the area in which the user can manipulate the query elements and the resulting elements. The result space is located at the surroundings of query space.

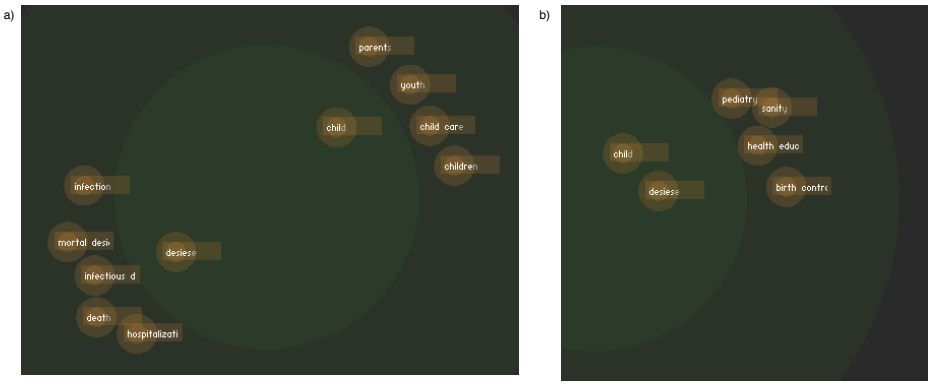


Fig. 4. (a) The results are given as a single query. (b) The keyphrase "child" and "disease" are separated, creating a separate query for each one.

Since the query statement is formulated by the amount of elements in a query space, the result will change with the position of each object. The amount of elements in a query space is considered. The right balance among the number of elements and the distance between them affects the return of the results.

6 Conclusion

In search tools, the problem of having long lists of results can be solved by considering the semantic relationships among keyphrases. With this option, the results of a search would be given, not as primer attention to the documents, but rather to keyphrases, these can be understood by the user, and also related. By learning from these relationships, the user can select relevant documents that match his/her search.

References

1. Beale, R.: Supporting serendipity: Using ambient intelligence to augment user exploration for data mining and web browsing. Int. J. Hum.-Comput. Stud. 65(5), 421–433 (2007)
2. Shah, U., Finin, T., Joshi, A.: Information retrieval on the semantic web. In: CIKM 2002: Proceedings of the eleventh international conference on Information and knowledge management, pp. 461–468. ACM, New York (2002)
3. Jones, S., Paynter, G.W.: Automatic extraction of document keyphrases for use in digital libraries: evaluation and applications. J. Am. Soc. Inf. Sci. Technol. 53(8), 653–677 (2002)
4. Kiryakov, A., Popov, B., Terziev, I., Manov, D., Ognyanoff, D.: Semantic annotation, indexing, and retrieval. Web Semantics: Science, Services and Agents on the World Wide Web 2(1), 49–79 (2004)
5. Wienhofen, L.W.M.: Using graphically represented ontologies for searching content on the semantic web. iv 00, 801–806 (2004)
6. Bhogal, J., Macfarlane, A., Smith, P.: A review of ontology based query expansion. Information Processing & Management 43(4), 866–886 (2007)
7. Jones, S., Paynter, G.W.: Human evaluation of kea, an automatic keyphrasing system. In: JCDL 2001: Proceedings of the 1st ACM/IEEE-CS joint conference on Digital libraries, pp. 148–156. ACM, New York (2001)
8. Khan, L., McLeod, D., Hovy, E.: Retrieval effectiveness of an ontology-based model for information selection. The VLDB Journal 13(1), 71–85 (2004)
9. Witten, I.H., Paynter, G.W., Frank, E., Gutwin, C., Nevill-Manning, C.G.: Kea: practical automatic keyphrase extraction. In: DL 1999: Proceedings of the fourth ACM conference on Digital libraries, pp. 254–255. ACM, New York (1999)
10. Katifori, A., Halatsis, C., Lepouras, G., Vassilakis, C., Giannopoulou, E.: Ontology visualization methods-a survey. ACM Comput. Surv. 39(4), 10 (2007)
11. Tzitzikas, Y., Hainaut, J.L.: On the visualization of large-sized ontologies. In: AVI 2006: Proceedings of the working conference on Advanced visual interfaces, pp. 99–102. ACM, New York (2006)

12. Jones, S., Staveley, M.S.: Phrasier: a system for interactive document retrieval using keyphrases. In: SIGIR 1999: Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, pp. 160–167. ACM, New York (1999)
13. Crestani, F., Vegas, J., de la Fuente, P.: A graphical user interface for the retrieval of hierarchically structured documents. Information Processing & Management 40(2), 269–289 (2004)
14. Corby, O., Dieng-Kuntz, R., Gandon, F., Faron-Zucker, C.: Searching the semantic web: Approximate query processing based on ontologies. IEEE Intelligent Systems 21(01), 20–27 (2006)

Modification of Web Content According to the User Requirements

Pavel Ocenasek

Brno University of Technology, FIT, Bozetechova 2, 612 66 Brno, Czech Republic
ocenaspa@fit.vutbr.cz

Abstract. The paper deals with the system for modification of web content according to the user requirements. The system is based on the network proxy server. The general idea is to employ translation rules (regular expressions) to render web pages upon user requests stored in profiles. The concept of the proposed proxy is general and the system can be used for various purposes. The main targets of use are: providing secure content of web sites, translation of non-accessible web pages into accessible form, etc.

Keywords: web content, translation, proxy server, security, accessibility.

1 Introduction

When the World Wide Web service was created and the markup language HTML became its main pillar of strength, only some people could foresee that it becomes one of the most valuable research or work instruments of wide society. Some of the best qualities that this service offers are availability and immediate diffusion of information published on the Internet. These characteristics are especially useful for users with some types of disability. Moreover, they have seen how their access to leisure, education, business or research activities has been improved.

2 Web Accessibility

To develop accessibility standards for Web sites and authoring tools, the W3C Consortium (www.w3.org) [2] [7] adopted the Web Accessibility Initiative (WAI). WAI guidelines group checkpoints into three levels of priority. Priority one includes checkpoints that Web site administrators "must" implement. For example, users must be able to avoid behavior that obscures the page content or disorients them. Flashing content can cause seizures in people with photosensitive epilepsy or distract cognitively impaired people. Distracting background images or sounds can affect those with visual or hearing problems. Priorities two and three are checkpoints that "should" or "may" be implemented [4] [6].

To avoid these problems, users must be able to filter WWW content or multimedia presentations. However, structure and meta information is hard to recognize and to filter. The main problems are:

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 324–327, 2008.

- to recognize and find titles
- to recognize and find links
- to recognize and find non-textual elements (such as inline images)
- to navigate from title to title
- to navigate from link to link.
- to handle input elements (such as entry fields, radio-, check- and other buttons)

3 System Concept

We have developed a new system, which will be useful for accessing web pages by visually impaired users and translate these pages into the accessible form. The system has been designed to make the web pages accessible independently [5] from the presentation devices and technologies used.

The main idea of the system can be seen from the following figure:

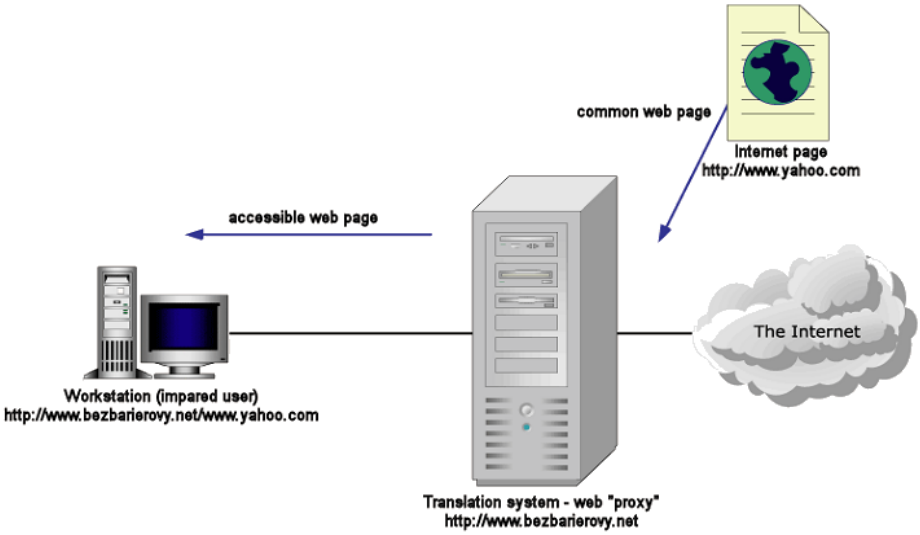


Fig. 1. The principle of automatic translation system. The system can be used either as a network proxy server (via proxy settings) or simple document proxy server (via URL prefix).

The system works as a proxy server for translating common internet pages into the accessible form. The web accessibility is described by translation rules, that are applied to the common pages.

The usage of our system is very easy. Before the first use, visually impaired user creates a profile where impairment-specific requirements for the translation are specified. Then the system is used via the standard web browser by specifying the URL to translate in the form: `http://www.bezbarierovy.net/www.yahoo.com` . The translation of the main page as well as all the linked pages that user visits from the starting page is done automatically.

4 Implementation

In general, the accessibility is performed according to the following instructions:

1. IMAGES – images could be easily switched off, resized or the color depth/contrast can be changed according to the user-specific requirements.
2. LINKS – visitors to the web pages are looking for information, and the more efficiently they can find it, the more valuable the site is to them. Most screen readers have a shortcut command that will give users a list of all the links on a page. This is a way to skim a page quickly.
3. COLOR – Consistent use of color can enhance the usability of your pages for many people. We have to be sure that no information is conveyed solely through the use of color.
4. TABLES – there are two simple things we can do to make tables more accessible without changing their appearance. One is to use the summary attribute. This attribute goes in the table tag along with the border, cell spacing and other attributes. The other thing we can do is to use the scope attribute in the first cell in each row and first cell in each column.
5. HEADINGS – those of us who are sighted use headings as a quick way to scan the organization of a page. To create headings, many people use the font tag to make larger text. However, most screen readers have a shortcut command that produces a list of all the headings on a page created with the heading tag. If the page is well organized and uses heading tags for headings, this can be a great way for visitors using screen readers to skim the page.

There are many rules and specific translations that belong to these (and other) categories. The detailed description is beyond the scope of this paper.

The proxy server can be used in two modes:

- Document proxy server, this mode is used when the impaired user enters the URL address in the standard browser in the following form: `http://www.bezbarierovy.net/<URL_to_translate>`. The system translates the starting page and automatically follows all links into the recursive translation.
- Network proxy server mode serves on a specified TCP port and translates all the content going through. The proxy server is activated by setting the proper address and port in the browser settings (the Connection/Proxy parameter). Then the common form of URL address is typed into the browser and the content is automatically translated.

In both modes of use the proxy server is transparent and browser independent. The translation is done according to the settings from the user profile.

5 Conclusions and Future Work

In this paper we have presented several tools that help visually impaired users to solve problems they lexperience when accessing information published on the Internet. Some of these problems can be analyzed from the Web designer's standpoint and the others from the user's perspective.

The main contribution of this paper is the presentation of the system, which is based on document-proxy techniques and translates web pages into the accessible form upon specified translation rules. The main advantage of the presented system is the universality of use and browser independency. Therefore, visually impaired users can use this system from various places with access to the Internet, such as home computers, libraries, school laboratories etc. Additionally, users can use their own stored profiles to make the browsing and accessibility more specific to their requirements.

Our next plan is to improve the user interface, the user-specific profiles and to simplify the rules into regular expressions, which will be used for translation of web content. We then will try to put these improvements into practical use.

Acknowledgement

The research has been supported by the Czech Ministry of Education in frame of the Research Intention MSM 0021630528: Security-Oriented Research in Information Technology, MSM 0021630503 MIKROSYN: New Trends in Microelectronic Systems and Nanotechnologies, and by the Grant Agency of the Czech Republic through the grant GACR 102/08/0429: Safety and security of networked embedded system applications.

References

1. Isaak J.: Toward Equal Web Access for All. IT Pro 11-12 / ’00, IEEE 1520-9202/00, 49-51 (2000)
2. Kirchner, M.: Evaluation, Repair, and Transformation of Web Pages for Web Content Accessibility. Review of Some Available Tools. In: Proceedings of the Fourth International Workshop on Web Site Evolution WSE 2002, 0-7695-1804-4/02 (2002)
3. Liu, S., Ma, W., Schalow, D., Spruill, K.: Improving Web Access for Visually Impaired Users. IT Pro 7-8 / ’04, IEEE 1520-9202/04, 28-33 (2004)
4. Macías, M., Sanchéz, F.: Improving Web Accessibility for Visually Handicapped People Using KAI. In: Proceedings of the 3rd International Workshop on Web Site Evolution WSE 2001 0-7695-1399-9/01 (2001)
5. Ocenasek, P., Toufarova, J.: Web Accessibility for Visually Handicapped People. In: INFORUM 2005: 11th Annual Conference on Professional Information Resources, Praha (2005) ISSN 1801-2213
6. Whitelaw, K.: Why Make Websites Accessibile? And How? In: SIGUCCS 2003, San Antonio, Texas, USA, pp. 259–261. ACM, New York (2003)
7. W3C World Wide Web Consortium. Cited 2008-01-30 (2006), `http://www.w3.org`
8. Ocenasek, P.: Automatic System for Making Web Content Accessible for Visually Impaired Users. In: WSEAS Transactions on Computers Research, Athens, GR, vol. 1(2), pp. 325–328 (2006) ISSN 1991-8755
9. Ocenasek, P.: Automatic System for Making Web Content Accessible for Visually Impaired Users. In: Proceedings of the 6th International Conference on Applied Computer Science, Puerto De La Cruz, ES, pp. 430–433 (2006) ISBN 960-8457-57-2

Virtual Environments with Content Sharing

Madjid Merabti[1], Abdennour El Rhalibi[1], Amjad Shaheed[1], Paul Fergus[1], and Marc Price[2]

[1] School of Computing and Mathematical Sciences,
Liverpool John Moores University,
Byrom Street, L3 3AF, UK
{M.Merabti,A.Elrhalibi,A.Shaheed,P.Fergus}@ljmu.ac.uk
[2] BBC Research, Kingswood Warren, Tadworth Surrey KT20 6NP, UK
Marc.Price@bbc.co.uk

Abstract. Content sharing over networked devices, beyond simple file sharing is becoming a reality. Many devices are forming closer relationships with different virtual worlds, such as World of Warcarft and Second Life. In one sense the gap between the two is becoming increasingly more blurred. Consequently, this opens up many new avenues for content sharing, not only between devices but also between sophisticated virtual worlds. Given such interoperable platforms a natural progression sees content that seamlessly resides within either. This will open up new opportunities where third-party content providers and users alike are able to create and share content over these new platforms. We aim to provide a basis on which this vision can be realised where mechanisms have been developed that facilitates the sharing of virtual world objects across different virtual environment. The work has been tested using a working prototype that allows digital content, to be shared and physical devices, such as mobile phones, to be connected and their content to be shared.

Keywords: Networked Virtual Environment, Content Sharing, Game Engine.

1 Introduction

Many devices have developed rapidly to become multifunctional wonders. They provide more functionality than they were originally designed to perform. The best example is that of the mobile phone. They not only provide communication functions such as making a phone call and sending text messages, but also work as a camera, MP3 player and support web access. More and more devices are providing computing capabilities, including increased networking functions enabling them to interact with each other more easily. The multifunctional capabilities of devices have given birth to exciting new application areas for networked appliances. Using the network, devices can be controlled from anywhere in the world. In online games such as *World of Warcraft* [1] and *Second Life* [2] players share a virtual environment in order to communicate, do business, and develop digital objects, which not only involve personal computer but also games consoles and mobile devices. Although real devices are used and the user is the only physical entity, these players communicate over large distance from different geographical locations from all over the world. Users can

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 328–342, 2008.

generate and share content, and buy and sell it in virtual environment. The challenge is to a achieve this within and across different heterogeneous virtual environments.

Many ways of distributing content may be possible; for example, between two physical devices; physical devices and virtual environments; or between different virtual environments and gaming platforms. Consider a scenario where a mobile user could share an audio file with a *Second Life* player. One approach would be to allow the user to simply ring the player's virtual mobile phone and send the track. Another example may be where a game user shares his or her assets and resources during a game with other potentially different games. This could be done for free or for a small fee. Here we could see a player in a massive multiplayer online game requesting and using resources, such as a weapon, or life time from some other player, in order to remain in the game longer. Here the user benefits from finding and using resources for free or at a very low price. Conversely, gaming studios benefit because more and more players will play or join the game because of the increased opportunity to extend game play and even earn money brings.

In this paper we propose a distributed framework based on a service utilisation framework which facilitates the sharing of content in virtual environments [3]. Using this framework networked appliances are automatically created and connected with associated avatars within a virtual environment. Our approach has many benefits, which include the user's ability to share, distribute and sell their content in virtual environments and the ability to remove the physical constraints associated with real world objects.

The remainder of the paper is structured as follows. In section II we introduce the background and the related work. Section III provides an overview of our proposed framework before describing a case study and a technical description of our approach in Section IV In Section V we provide our conclusions and future work.

2 Background and Related Work

Massively multiplayer online games already attract huge numbers of players and are expected to become increasingly popular where they are already forming the basis for next-generation gaming. Utilising Internet communications, games have blurred virtual and physical worlds and converged with social networks [4]. This has changed how users view and play games. Many games such as *Planetside* [5], *Star Wars Galaxies* [6], *The Sims Online* [7] and *EVE Online* [8], are dependent on network communications. None more so than the game *World of Warcraft*, which became the fastest selling PC game in North America in 2004-2005 and in 2006 was reported to have 6 million subscribers worldwide [1].

Although multiplayer gaming clearly provides significant benefits over single-player games through the use of networking, its client-server architecture enforces a number of limitations. Most notably, game play and enhancements must be carefully controlled through centralised gaming servers. This results in bottlenecks, central points of failure, and the inability to appropriately react to real-time changes in large virtual worlds. Gamers are tied to games through proprietary software and hardware installations. User interactions do not affect strategic developments and games do not support self-management capabilities to extend functionality beyond those they have been pre-programmed with.

This has lead to shifts within the gaming industry, where increasing access to game engines, software development kits and level editors has allowed games to be changed more easily. This phenomenon – known as modding – marginally alleviates some of the limitations discussed above [9-11]. Although modding provides a means of adapting and evolving games, it is restricted to more technically savvy users, such as software developers, rather than people who simply just play games. Furthermore, mods are tied to specific games. For example, a mod developed for the *unreal engine* will be incompatible with the *quake* engine. Some researchers suggest that distributed technologies in conjunction with middleware may relieve many of these difficulties, however it is generally accepted that more research is required to establish a suitable architecture [12].

Modding is an activity that runs alongside mainstream games development, with developers providing modding tools as a way to attract customers. In essence modding is seen as a business strategy. Although not explicitly stated, incentives to mod games are used as a means of generating free development for publishers, for example through the use of modding competitions that act as a means of screening game enhancements in order to include them in future releases. In most cases this is an unpaid source of labour and gaming organisations carefully control how it is executed [9]. Through competitions and gaming subscriptions for massively multiplayer online games, the industry has a healthy flow of mod software. In support of this several game companies adopt the principle of modding as a key strategy, where only a base solution is initially provided. Any enhancement to the game thereafter is dependent on user modifications. One example of this is BioWare's *Neverwinter Nights,* which is heavily reliant on gamer-created content [13]. Successful mods have been incorporated into subsequent releases. Another example is *Counterstrike,* which is a modification for team play of Valve Software's *Half-Life* [13]. In this case modding can be seen as an important and welcome source of innovation where commercial risks are not taken by the gaming industry, but rely on the goodwill of the modders [10].

Whilst modding has been discussed from a game coding perspective, mods may also exist as part of and within the game itself. Communities such as *Second Life* [14] are heavily reliant on users shaping the virtual environment, extending the concept of MUDs into realistically rendered virtual worlds [15]. Graphical objects of any description can be developed and added to the virtual world, which can then be shared or sold between avatars' within that world. Modifications to the environment (e.g. land) can be made and buildings can be constructed. This differs somewhat from conventional modding in that all modifications take place within the virtual world. However, there is no mechanism to allow the objects created in *Second Life* to be shared and distributed amongst different online games and a better approach could be used to expose these modifications so that they can be utilised universally.

The increasing popularity amongst multiplayer gaming platforms shows that they are being used more than for just passing time. The virtual world's platform is already being utilised for business but it may also be used for community and financial analysis of the gaming business itself. Among other uses we see them being used for specialized training for armed forces or vocational training, medical consultation and psychoanalysis, and for community and financial experimentation to analyse social norms [16].

The study shown in paper [17], plans to build a virtual world where virtual objects visualize the information collected in wireless sensor networks by which virtual worlds may allow peers to understand and estimate more easily the state of the real world measured through sensor networks.

Many of us now have access to home technologies such as computers, game consoles, and the internet which can help us to create content. There is however limitations with current approaches where they fail to investigate how young people participate in content creation or what tools they use and the extent of their commitment. Such surveys are required to better identify the potential of content creation [18].

The research carried out in [6] proposes to examine cognitive overload problems among game players who always interact with the game world as well as with other users [6]. Using Maple Story [19] as a case study; the authors have found different results which show different types of cognitive overloads to come into sight during game play which might cause serious problems to all players.

This paper [20] introduces Konark, a service discovery and delivery protocol for ad-hoc, peer-to-peer networks in which the authors provide an infrastructure to set up generic peer-topeer systems. It acquires advantages of basic networks for peer naming and message routing. It uses entirely distributed, peer-to-peer techniques for resource discovery which provides every peer with the ability to publicise and discover the resources in ad-hoc networks [20].

There are many business opportunities and challenges involved in the virtual world where millions of people from all over participate in the network to play online games. This paper

[21] particularly discusses the principles and policies related to the social implications of *Second Life* [22] which raises significant research questions. One of the important questions, for example, is the payment issue either to the avatar or customer, while another issue is taxing people who are earning money in the virtual business. As can be seen there is considerable research interest around virtual environments with each approach making some very interesting contributions. In the remainder of this paper we discuss our proposed framework that builds on these advances to extend current approaches where user generated content can be more freely shared across different virtual environments.

3 Framework Overview

Our approach is based upon the service utilisation framework proposed in [3] and depicted in Figure 1, we have designed a plug-in containing several services that allow users to share their resources.

This component provides an interface between applications such as gaming platforms, virtual worlds and network devices. The *Visual Resource Manager*, as is illustrated in Figure 2 extends the functionalities provided by the service utilisation framework.

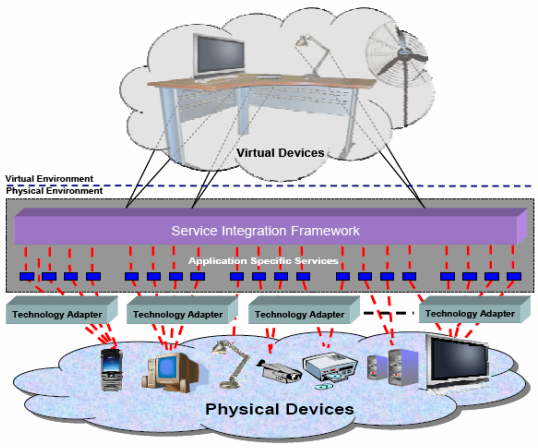


Fig. 1. Service Utilisation Framework

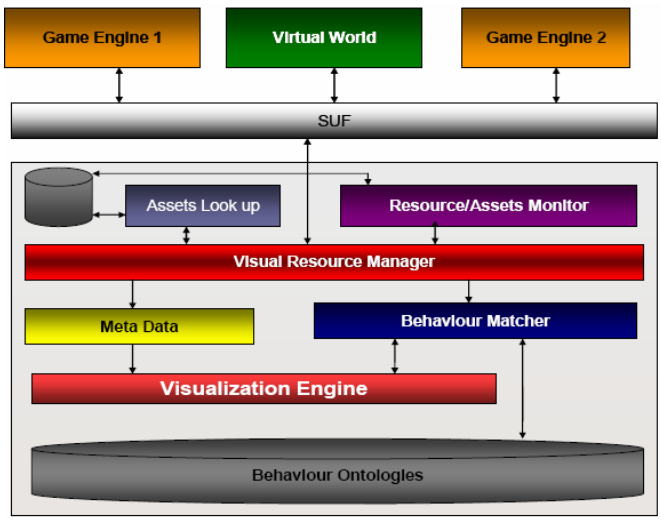


Fig. 2. Component Diagram

This component consists of different services such as the *Resource Monitor*, *Resource Lookup*, a *Meta data engine*, a *Behaviours Matcher*, and a *Visualization Engine*. Using these services along with those provided by the service utilisation framework physical resources are linked with their digital counterparts that reside in the virtual world.

Requests received from users are matched against entries in the *Lookup Service* and the *Resource Monitor* is used to monitor interactions between these resources. The communication packets used in the framework are serialised as XML. XML enables the sharing of structural data across different formats especially through the Internet. It

also allows descriptions to be extended through the addition of new tags which makes it efficient. Ontologies, a shared understanding of some domain, are used to promote better understanding of relationship between the same concepts using different terminologies.

When we receive an object the Virtual *Resource Manager* registers the resource with the resource *Lookup Service*. Following this it extracts the meta-data of that object and passes it to the *Visualization Engine* which in turn renders the 3D object into a graphical shape. At the same time the Behaviour Matcher lookups up the behaviour of similar objects in that environment. If it finds the behaviour then it assigns it to the object along with the effects it supports when it is executed. Let us assume that a user sends a game object to another game. The game objects behaviours should also be transferred from the source environment to the destination environment so that the user can fully enjoy the new object features such as graphical special effects or how it reacts to stimulus from the game, such as being shot at.

The object consists of two layers; the first layer contains Meta data used to describe the object as illustrated in Figure 3, which is an XML file containing different attributes of that object as shown in Figure 8 below.

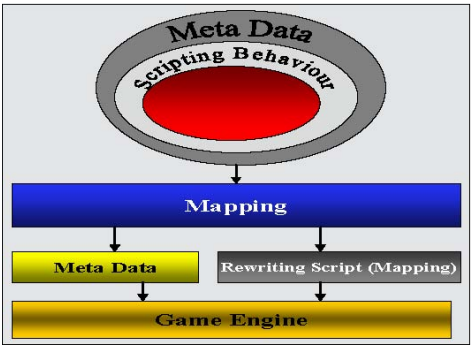


Fig. 3. User Generated Object

The second layer is the scripted behaviours. A mapping is performed between the object and the game engine by extracting the meta data used to describe the object and its 3D characteristics and the rewriting scripting engine used to find appropriate behaviours the game engine can accommodate as detailed in the scripted behaviour of the object.

Using the principles of service-oriented computing, components, such as game consoles and mobile phones, implement a small footprint of code allowing functions, such as audio and video, to be disseminated within the network. Using the framework services, the components can link to the network using any communication protocol; discover and/or publish and use framework and application services locally (provided by the component itself) or remotely (provided by other components); carry out semantic interoperability between different vocabularies used by component manufacturers; automatically form communication links with other components in the network; self-manage links with other components in the network; and self-manage their operations based on composite and environmental changes. Application particular

services, on the other hand offer a means of dispersing and utilising component functionality (such as audio and video), gaming engines, and player (AI behaviours) and game objects (tree, car or avatar).

This is achieved using the service integration framework [13], implemented on every component – be it a networked appliance or software module from the virtual world. This is a peer-to-peer interface that can be mapped onto any middleware model. Devices connect to the network as either specialised components or simple components. A specialised component has the ability to provide services as well as to propagate service requests within the network. A simple component by comparison has more restricted abilities: it joins the network, propagates queries and invokes discovered services. For example, this could be sensors in a network that provide multimedia data for crowds or flocking. This enables any component irrespective of its capabilities to effectively choose how it will interact within the network.

Using this architecture, we have designed a distributed service-oriented platform for use with virtual environments and physical devices. This allows multimedia content to be shared with the virtual environment from any physical multimedia producing device such as mobile which we discuss in more detail below.

Whilst it is important to bear in mind the overall structure that a virtual environment might take, it has been a goal of our work to deconstruct as far as possible the holistic notion of a virtual environment into a set of autonomous, generalised and reusable components. Whilst the development process of our framework necessarily entailed the compartmentalisation of various aspects of a traditional game, the final result must therefore be considered from the opposite perspective. Ultimately we aim to allow gaming to exist as an ad hoc interaction between various networked components, the entirety of which forms the virtual environment. None of these components in isolation can be considered to be the virtual environment itself. Perhaps the closes to what might be considered the heart of the virtual environment might be the rendering or physics engines. However, these will only provide one of any number of interpretations of the interactions that occur between components.

3.1 Behaviour Ontologies

Ontologies allow communication and a common understanding among game objects from different gaming environments that have never seen before nor even heard of. Thus, new game objects may be introduced to the game at any moment and be accepted by the already-playing ones. By using ontologies we can define object properties and also hierarchical service interfaces (for game-object communications). Ontology-aided design may also be helpful at the game planning stage to design the whole game universe (the game-object-related classes and the game-objects themselves). Developed ontologies might then be very easily incorporated into game-objects. To give an example of a simple ontology, we can say that bullet belongs to the class of ammunition which are both affectable and can affect. Affectable means that another game-object may change the bullet's properties (for a while or even constantly: e.g. shoot it and thus make it disappear). Using this, a player can shoot an opponent - shortening his life. Can affect would mean that the game-object can influence other game-objects. With ontologies we can make use of such complex hierarchies and relationships in a simple way.

In terms of behaviour, we consider our game-objects as characters. There are many models we can use to predefine a game-object's behaviour. For the most complex example we are interested in techniques for which the character's behaviour is not completely determined in advance. To determine the behaviour we apply reactive behaviour rules. The use of reactive behaviour rules was one of the first approaches proposed for generating character behaviours, and it is still one of the most popular and commonplace techniques. Great success has been obtained in developing rule sets for various kinds of behaviour, such as flocking and collision avoidance. As an example of a simple stimulus-response rule that can result in extremely sophisticated behaviour.

3.2 Behaviour Matcher

Existing work on game object/character behaviour modelling can be generally classified into a microscopic approach and macroscopic approach. Most computational models for object/character modelling and simulation adopt the microscopic approach where each individual agent is equipped with a set of decision rules to determine what to do in the next time step. The object/character behaviours are then naturally generated as some emergent phenomena due to the interactions of the individual object/character.

In our system, a two-level cognitive model architecture is adopted. The lower level is used to model individual behaviours, and the top level model is used to represent object/character dynamics and interaction. This two-level architecture is a natural reflection of the interaction amongst object, and between an object and a device in real-life situations. An interaction can emerge amongst individuals and might take into account environmental factors.

Individuals involved in this emerging process may change their behaviours after an interaction is formed. When an object/character joins the new environment, the behaviour of the individual in the new environment will be determined by both the environment model and the object/character behaviour model.

n our system, the Protégé (http://protege.stanford.edu) ontology knowledge repository is used with the JESS inference engine to keep track of the environment and the behaviours of objects/characters in the system. The execution environment will provide updates on changes in both the environment as well as the status of object/characters and human players. These changes will be updated into the knowledge-base and the inference engine will modify the behaviours of individuals accordingly based on the cognitive model.

4 Homura Game Engine and IDE

The game engine and IDE used is Homura. The initial architecture of Homura as a whole is to have a core engine, which uses jMonkeyEngine and LWJGL. The user is able to create so-called Homura projects which will run totally independently of the IDE and can be exported to a wide range of platforms. Homura IDE is a powerful IDE that is based on the Eclipse Platform, and uses existing Eclipse plugins and technology. One particularly important plugin the IDE uses is the JDT (Java development Tools), which provides the user with a rich Java editing environment for creating their game logic in. Figure 4 shows the interface and a game example developed with this platform and Java Monkey Engine (jME).

Various parts of Homura are declaratively specified in XML files in the root of a Homura project, and this provides a link between the classes and concepts used in a Homura project. These XML files come in useful when considering exporting to a website, as the website can parse these files in a standard way, and act upon the data contained within them. The IDE can also act upon, and manipulate these files to change various parts of a Homura project. This is similar to how Eclipse works with the plugin XML that sits in the root of each RCP project, allowing concepts to be linked to classes, and functionality exposed to other plugins.

The IDE itself hooks into the running Homura engine while a Homura application or game is running to provide various introspection and debugging facilities. For example, the user is able to see details about the concepts which are in operation within the application, as well as the current frame-rate through a statistics view. One feature that is particularly important for the user to inspect is the scene graph, as this allows them to find out why their graph is not correct and help them find the area of code which is manipulating it incorrectly. They are also able to view the various properties of the scene elements they can select, for example, the world and local translation of a node.

In order to provide the necessary hooks into the running engine, some parts of jME have been modified, as modifying Homura alone may not be enough, or too inefficient at certain levels. For instance, it may be difficult to tell when a scene graph has changed if it is being modified programmatically through the user's own code. Other parts of the engine can be probed at intervals to check their status.

Parts of the Homura IDE use Homura and jME, not just the games and applications that the user creates. However, if Homura is providing a game interface where parts of the API are accessed in a game-like context, the method of integrating parts of this engine into the IDE will be less than ideal. At the moment, jMonkeyEngine has a game-specific context. The idea is to unravel this into a hierarchy of non-game specific classes and interfaces, with the game-specific classes and interfaces at the top of this, with the notion of being able to run an 'application' and a 'game'. Then, only the game-specific classes will have access to game-specific concepts. Therefore, it will be necessary to keep the game-specific details separate from the application-specific ones.

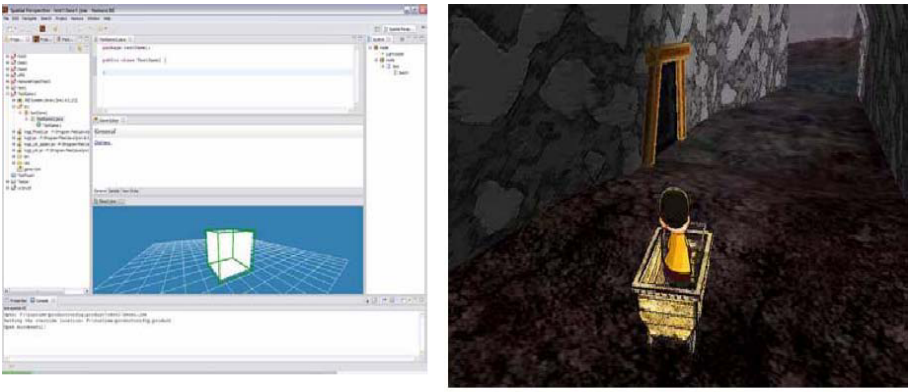


Fig. 4. Homura Interface and Game Application

5 Case Study

A case study has been conducted to demonstrate our approach that shows one way of sharing content. We have developed our scenes using blender. For example, Figure 8 shows and XML representation of a gun. In order to load these XML files into the scene graph of jME the function illustrated in Figure 5 is used.

```
1  private Node LoadXML(String filePath){
2      try{
3          //  Initialise binary converter
4          XMLtoBinary converter = new XMLtoBinary();
5
6          //Initialise binary reader
7          JmeBinaryReader jbr = new JmeBinaryReader();
8
9          // convert the input file to a jme-binary "LabCube"
10         ByteArrayOutputStream LabNoExtra = new ByteArrayOutputStream();
11         URL LabModel = ModelLoader.class.getClassLoader().getResource(filePath);
12         converter.sendXMLtoBinary(new BufferedInputStream(LabModel.openStream()), LabNoExtra);
13
14         //get the "LabCube"
15         tempNode = new Node ("Temporary Node");
16         tempNode = jbr.loadBinaryFormat(new ByteArrayInputStream(LabNoExtra.toByteArray()));
17     }
18     catch(IOException e){
19         System.out.println("Couldn't load the input file:" + e);
20         e.printStackTrace();}
21     catch (java.lang.NullPointerException npe ) {
22         npe.printStackTrace();
23     }
24     return tempNode;
25 }
```

Fig. 5. Loading XML Serialised Scenes

jME doesn't support loading any file formats directly. Rather it uses jME binary, its own format. Different classes included in jME convert scripts are used to create the jME binaries. First of all the binary converter and binary reader is loaded as it is illustrated in Figure 5 from lines 4 to 7. As illustrated in line 10 and 12 the OutputStream and the InputStream are used to send and read the appropriate contents. The XML file is converted with the ByteArrayOutputStream and read with the ByteArrayInputStream. This process allows us to transfer meta data from one environment to another environment as shown in figure 6.

Using our framework we have implemented an application that links a mobile phone to a corresponding avatar in the jME as depicted in Figure 7. Through this connection we are able to us the physical phone and its avatar representation in the virtual world. We can answer and make calls from the physical and the virtual and using the same communication channels we are able to transfer user generated content between the two. For example, we can pass a music track to the virtual world along with metadata describing what it looks like and the behaviours it supports.

In Figure 7 we see a mobile phone and for each song passed from the physical device to the virtual mobile phone a radio button is added and visually connected to the avatar phone. When a song is selected the behaviours appears in the virtual world; in this case we can see that by selecting Song1 we can execute one of its behaviours, i.e. exit, play, and stop, by pressing the buttons located to the right of the virtual world screen. In the following section we discuss the technical aspects in achieving this.

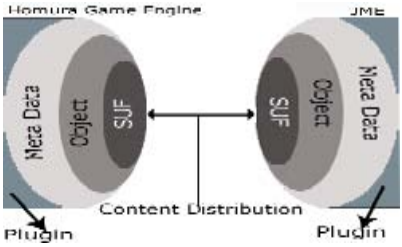


Fig. 6. Content Sharing

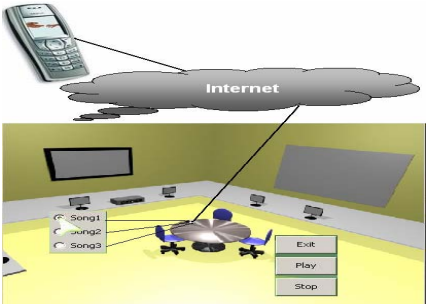
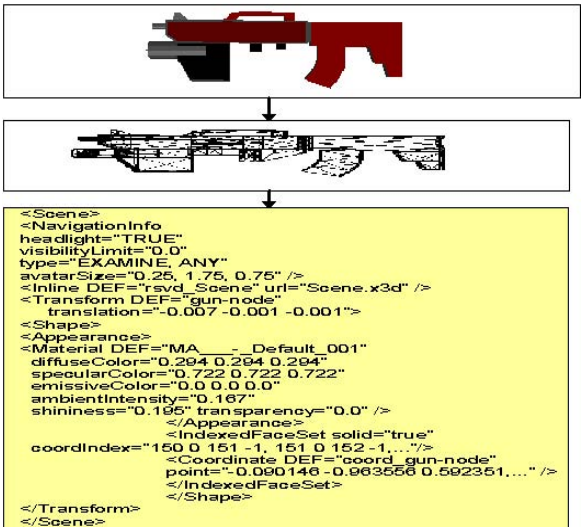


Fig. 7. Virtual world

6 Technical Description

In the architecture described above we have designed a distributed service-oriented platform to link between networked appliances and associated avatars in virtual environments we have been able to carry out experiments using our design and show how multimedia and gaming content can be shared inside virtual environments. Using JXTA [23] as its peer-to-peer middleware protocol, a virtual environment developed using jME [24] queries the network for JXTA services advertised by the peers (Physical mobile phone). We have connected two virtual environments using JXTA, one game developed in Homura Game Engine and the other our virtual lab, developed using JME.



```
<Scene>
<NavigationInfo
headlight="TRUE"
visibilityLimit="0.0"
type="EXAMINE, ANY"
avatarSize="0.25, 1.75, 0.75" />
<Inline DEF="rsvd_Scene" url="Scene.x3d" />
<Transform DEF="gun-node"
    translation="-0.007 -0.001 -0.001">
<Shape>
<Appearance>
<Material DEF="MA____-__Default_001"
  diffuseColor="0.294 0.294 0.294"
  specularColor="0.722 0.722 0.722"
  emissiveColor="0.0 0.0 0.0"
  ambientIntensity="0.167"
  shininess="0.196" transparency="0.0" />
                    </Appearance>
                    <IndexedFaceSet solid="true"
  coordIndex="150 0 151 -1, 151 0 152 -1,...">
                    <Coordinate DEF="coord_gun-node"
                    point="-0.090146 -0.963556 0.592351,..." />
                    </IndexedFaceSet>
                    </Shape>
</Transform>
</Scene>
```

Fig. 8. Meta Data describing the rendering information for a gun

In the above case study a peer makes a request for a service, such as a game object (e.g. gun), in the virtual lab where another peer has previously advertised its sharable assets using JXTA advertising services. In Figure 8 we show, how we have implemented the scenario in which the user requests a gun resource.

We pass the meta-data to jME, which in turn is used to render a 3D representation of the gun in the scene. The gun object also contains the scripting behaviours it supports. For example, Figure 9 illustrates, in part a simple script for the fire behaviour. Javascript was used and where developed using the Rhino API from Mozilla [25], which is used with the Java Scripting Framework [26] and the open-content repository API provided by Captive Imagination [27].

```
var incr = 1.0;

function fire() {
   if (ammunition > 0) {
      if(consoleI=undefined)
         console.log("time="+time + "Fired single round");
      else
         System.out.println("time="+time + "Fired single round);
         //update the ammunition
      ammunition = ammunition - incr;
      controlled.updateGunFired();

   }
}
```

Fig. 9. Sample behaviour

```
import com.sample.DroolsTest.Message;

rule "Fire"
   when
      m : Message( status == Message.FIRE, message : message )
   then
      System.out.println( message );
      m.setPayload("function fire(){if (ammunition > 0)
         {if(consoleI=undefined)console.log('time='+time + 'Fired single round');else
         System.out.println('time='+time + 'Fired single round);ammunition = ammunition
         – incr;controlled.updateGunFired();}"
      m.setMessage( "Fired" );
      m.setStatus( Message.FIRED );
      update( m );
end
```

Fig. 10. Rules used to create scripted behaviour

The Meta data and scripts, including the aforementioned tools where applied in the same way to allow music to be shared between our mobile phone and its associated avatar in the virtual lab. The goal here is to show how two very different types of content can be shared. One associated with multimedia and the other associated with conventional game playing objects. Perhaps these act as two extremes between which many other possibilities are possible.

Both meta data for objects and the scripted behaviours are passed between different environments using JXTA pipe and messaging objects in which all required information is presented to extract and construct the associated object. Whilst, we simply use the meta data to construct the objects, we run all scripting behaviours through a set of rules, as discussed in the above section on the Behaviour Matcher. Whilst objects may support behaviours in their source environment, it is not necessarily the case in the target environment. Here the rules try to extract the behaviours the target environment supports. The Behaviour Ontology acts as an interoperability mechanisms between terminologies which we have implemented and serialised using the Web Ontology Language (OWL) [28]. The rules where developed using Drools, where Figure 9 shows in part a simple fire and part of the script for a behaviour being generated.

In the mobile phone scenario we demonstrated how two users are able to share multimedia content between physical mobile device and corresponding avatars in the virtual world. We stream multimedia content from the physical mobile device to the virtual mobile using the Java Media Framework (JMF) [29] and the Real-time Transmission Protocol (RTP) [30].

RTP packets were wrapped in JXTA [23] messaging objects to abstract the IP dependent format used for HTTP calls in RTP. This provides a unified addressing scheme ensuring that all components are addressed in a uniform way. Frames were sent from the physical mobile to the virtual environment using JXTA Pipes. Upon receiving the JXTA packets, the RTP packets are extracted and processes by a custom data source adapter developed for the purpose, which streams RTP data much as it is done traditionally, after network connectivity , the avatar requests the list of song in which he/she then chose a song from the list, after that the mobile start streaming the content using RTP protocol in JXTA pipes; when it receives first stream JMF process the stream and check supported codec, if it is supported then it will continue receiving streams from mobile device in our case while playing the stream using JMF and a plugin called Fobs4JMF [31] which supports most formats such as mp4 or 3gp. These tools in conjunction with Skype allowed us to enable bi-directional communications between the physical mobile phone and its virtual world counterpart.

7 Conclusions and Future Work

In this paper we presented a novel framework that allows content to be shared across different virtual environments. This extends current gaming platforms in a number of different ways which we have discussed in this paper. Our framework is based on a novel approach that draws on our expertise in the area of Networked Appliances and gaming. This provides a novel perspective on how virtual worlds can utilise the benefits from both to form a blurring that allows content to be easily shared and used from within both. Interpretations can be made about content shared, which may include conventional multimedia as well as well known digital content such as guns, and life for game play. This allows a more intimate link between heterogeneous games where such interpretations form the basis for visual renditions as well as the ad hoc generation of behaviours those objects being shared support. Whilst understanding the effects a car crash may have in a driving game, such interpretations allow the car to inhabit a world in which the concept of a car is not necessarily understood, but where behaviour mappings allow a comparison to be made between the effects of a crash and for example, that of being shot at. This not only makes games more flexible, but it also provides a basis for more interesting virtual environments not yet seen.

As well as benefiting the gaming community this multidisciplinary approach might provide additional functionality through interactions with real-world networked appliances, allowing facets of physical devices to be projected into the game in order to provide virtual manifestations of themselves. Whilst we have presented an initial prototype system, it is clear that much work remains to be carried out before a fully effective system is produced. In particular, working with rules and dynamic script writing in conformance with ontologies needs to be better understood. We hope to extend the use of ontologies in the system in order to increase the robustness of interactions between components, allowing for greater flexibility in the way components represent themselves.

This is a multidisciplinary project spreading across several research areas. The goal is to create a tighter relationship between the advances we have already made in create a new framework that incorporates all others, i.e. the service integration framework, the

content sharing services, the Homura engine and IDE and the Drools Rules and Java Scripting Framework. This will be the focus of much ongoing work. Ultimately, the success of a framework such as this relies on the development of exciting content that can be used to build up gaming environments. Nonetheless we believe that a flexible and distributed system such as this provides many opportunities for the advancement of gaming, virtual environments and networked devices in the physical world into new areas and in new ways.

References

[1] Ducheneaut, N., Yee, N., Nickell, E., Moore, R.: Building an MMO with Mass Appeal: A look at Gameplay in World of Warcraft. Games and Culture - Journal of Interactive Media 1(4) (2006)

[2] Herman, H., Coombe, R.J., Lewis, K.: Your Second Life? Cultural Studies 20(2-3) (2006)

[3] Merabti, M., Fergus, P., Abuelma'atti, O., Yu, H., Judice, C.: Managing Distributed Networked Appliances in Home Networks. Proceedings of the IEEE Journal (2008)

[4] Seay, A.F., Jerome, W.J., Lee, K.S., Kraut, R.E.: Project massive: a study of online gaming communities. In: CHI 2004 extended abstracts on Human factors in computing systems, pp. 1421–1424. ACM Press, Vienna (2004)

[5] PlanetSide (2006) (Accessed 2007), `http://planetside.station.sony.com/`

[6] Star Wars Galaxies (2006) (Accessed 2007), http://starwarsgalaxies.station.sony.com/

[7] The Sims Online (2006) (Accessed 2007), `http://www.ea.com/official/thesims/thesimsonline/`

[8] EVE Online (2006) (Accessed 2007), `http://www.eve-online.com/`

[9] Sotamma, O.: Have Fun Working with Our Product!: Critical Perspectives on Computer Game Mod Competitions. In: International DiGRA Conference, Vancouver, Canada (2005)

[10] Kucklich, J.: Precarious Playbour: Modders and the Digital Games Industry. International Journal on Fibreculture 1(5) (2005)

[11] El-Nasr, M.S., Smith, B.K.: Learning Through Game Modding. ACM Computers in Entertainment 4(1) (2006)

[12] Hsiao, T., Yuan, S.: Practical Middleware for Massively Multiplayer Online Games. IEEE Internet Computing 9(5), 47–54 (2005)

[13] Computer Game Modding, Intermediality and Participatory Culture, F. University of Tampere (2003) (Accessed: September 2003), `http://old.imv.au.dk/eng/academic/pdf_files/Sotamaa.pdf`

[14] Yee, N.: The Unbearable Likeness of Being Digital: The Persistence of Nonverbal Social Norms in Online Virtual Environments. The Journal on CyberPsychology and Behaviour (to appear, 2006)

[15] Curtis, P., Nichols, D.A.: MUDs grow up: social virtual reality in the real world. In: COMPCON 1994, pp. 193–200. IEEE Computer. Soc. Press, Los Alamitos (1994)

[16] Book, B.: Moving Beyond the Game: Social Virtual Worlds. Series Moving Beyond the Game: Social Virtual Worlds (2004)

[17] Kwon, T., Choi, S.-M.: Deriving the Virtual World from Wireless Sensor Networks for Interaction with Consumer Electronic Devices. Consumer Electronics. In: ICCE 2007 (2007)

[18] Hayes, E.: Game content next term creation and it proficiency: An exploratory study (2007)

[19] Maple Story (Accessed 2007), `http://maplestory.nexon.net/`

[20] Desai, N., Verma, V., Helal, S.: Infrastructure for Peer-to-Peer Applications in Ad-Hoc Networks (Accessed 2007), http://www.harris.cise.ufl.edu/projects/publications/konark_p2p.pdf
[21] Papagiannidis, S., Bourlakis, M., Li, F.: Making real money in virtual worlds: MMORPGs and emerging business opportunities, challenges and ethical implications in metaverses. Technological Forecasting and Social Change (2007)
[22] Second Life (Accessed 2007), http://www.secondlife.com
[23] JXTA (Accessed 2007), https://jxta.dev.java.net
[24] Java Monkey Engine User Guide (Accessed 2007), http://www.jmonkeyengine.com
[25] Rhino: JavaScript for Java. 2007, Mozilla.org (Accessed 2007), http://www.mozilla.org/rhino/
[26] O'Connor, J.: Scripting Framework for Java (2006) (Accessed 2007), http://eventhorizongames.com/wiki/doku.php?id=articles:java_scripting_framework
[27] The open-content repository for games. Captive Imagination (Accessed 2007), http://captiveimagination.com/svn/public/cigame/trunk/
[28] Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web, vol. 284(5). Scientific America (2001)
[29] Java Media Framework (JMF) (Accessed 2007), http://java.sun.com/products/java-media/jmf/
[30] RTP: A Transport Protocol for Real-Time Applications (Accessed 2007), http://www.ietf.org/rfc/rfc1889.txt
[31] RTP: A Transport Protocol for Real-Time Applications 1996, IETF (Accessed 2007), http://www.ietf.org/rfc/rfc1889.txt

Hand Contour Tracking Using Condensation and Partitioned Sampling

Daiguo Zhou, Yangsheng Wang, and Xiaolu Chen

Institute of Automation, Chinese Academy of Sciences,
Beijing, 100080, China
{daiguo.zhou,yangsheng.wang,xiaolu.chen}@ia.ac.cn

Abstract. In this paper, we present a visual articulated hand contour tracker which is capable of tracking in real-time the contour of an unadorned articulated hand with the palm approximately parallel to the camera's image plane. In our implementation, a B-spline deformable template is used to represent human hand contour, and a 14-dimensions non-linear state space which is divided into 7 parts is used to represent the dynamics of a hand contour. The tracking is performed in grey-scale skin-color image based on particle filter and partitioned sampling. Firstly, a Gaussian model is used to extract the skin pixels. Secondly, particles for each of the 7 parts of the non-linear state space are generated hierarchically based on second-order auto-regressive processes and partitioned sampling, and then each generated particle is weighted by an *observation density*. Finally, the best complete particle is chosen as the tracking result, and several complete particles are stored to be used in the next frame. The experiments show that our tracker performs well when tracking both rigid movements of the whole hand and non-rigid movements of each finger.

Keywords: hand tracking, condensation, particle filter, partitioned sampling.

1 Introduction

Visual articulated hand tracking is an attractive area of research in the computer vision community. It has a great potential in VR, AR, HCI and computer games. However, the high degree of freedom of hand configuration, the self-occlusion of fingers, and the kinematic singularities in the fingers' articulated motion, make visual articulated hand tracking become a difficult and challenging problem.

Because of the wide range of potential applications and the technical challenges, visual hand tracking has attracted a great deal of researches onto it during the last decade. In 1993, Rehg builds a system called DigitEyes [1] where a hand could be tracked against a black background. In the system, a kinematic 3D hand model whose initial configuration is known is used to represent human hand. The state of the model is updated by solving a constrained nonlinear optimization problem. In addition, self-occlusions of the fingers are handled using layered templates. Rehg's work is seminal since it establishes a classical approach for model based tracking of articulated hand,

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 343–352, 2008.

however, the adaptation of the kinematic 3D hand model to a new user is time-consuming, and finger occlusions are only tracked off-line in the system, what's worse, the optimization process may get trapped in local minima which could result in the tracking failed. In another system similar to Rehg's, Kuch and Huang [2] simplify the model adaptation to a new user with only three snapshots of the user's hand in three predefined configurations. Though making some improvements to the DigitEyes system, their new system still has the local minima problem. In 2001, Stenger et al. [3] construct an anatomically accurate hand model using truncated quadric, and an unscented kalman filter is used to estimate the pose of the hand. In 2006 [4], Stenger et al. improve their approach by first discretizing the state space and organizing it into a hierarchy of hand templates, and then searching down the template hierarchy to refine the fitting of the hand model to the hand in an image. Their method produces good results, and it is capable of handling out-of-image-plane rotations, fast motion, and automatic recovery of tracking. However, the system has large memory requirements and does not work in real-time.

In 1998, Blake and Isard [5] establish another important kind of hand tracking approach based on deformable 2D contours and particle filter. In the paper, they introduce the Condensation algorithm and give a number of experiments to demonstrate the robustness of their method. However, it's inefficient to track an articulated object which has a high dimension state space using Condensation alone. In 1999, MacCormick and Blake [6] introduce a new technique called partition sampling, which makes possible to track more than one object. A year later, MacCormick and Isard [7] implement a vision based articulated hand tracker using this technique. Their tracker is able to track position, rotation and scale of the user's hand while maintaining a pointing gesture. Based on Blake's work, Martin Tosas [8] implement a full articulated hand tracker which is then used for Virtual Touch Screen in his thesis. He makes some technique extensions to Blake and MacCormick's methods.

Besides the approaches mentioned above, the following are some other kinds of hand tracking methods. Nolker and Ritter's algorithm [9] firstly find the fingertips of a hand in a grey-scale image by means of a hierarchical neural network, and then update a 3D hand model from the fingertips positions. Stefanov [10] combines Hough transform features with behavior knowledge in order to guide and achieve robust hand tracking. Kolsch and Turk [11] use a technique called flocks of features to track a hand through rapid deformations against complicated background.

Our implementation is based on Martin's work [8], while some changes are made: the skin extracting method is simpler, the measurement operation is more effective, the calculation of the weight for a complete particle is new and reasonable, and instead of particle interpolation used in Martin Tosas [8], we use a simpler selection strategy to form complete particles.

The rest of the paper is organized as follows. Section 2 introduces briefly a simple but effective skin-color extracting method. Section 3 gives out a deformable hand contour template and a non-linear state space which represents the movements of the hand contour. In Section 4, we describe how to track in real-time the hand contour throughout a video sequence. Section 5 shows some experimental results. And Section 6 is a short summarization.

2 Extracting Skin Color

Our tracker is skin-color based. Before tracking performs, we need to extract skin-color from video. In this paper, we use a single 2D Gaussian on Cb-Cr space to model skin-color. The mean vector and covariance matrix of the Gaussian model are estimated in advance using some well-chosen training data. To extract skin color in a 24-bit image, some pre-process are performed at first, then each pixel is transformed into YCbCr space, and the CbCr sub-vector is tested against the 2D Gaussian model to decide whether the pixel belongs to skin color or not, only the pixels belong to skin color kept. Though the model is simple, the result it produces seems good and satisfies our application. Figure 1 shows the result of this method.

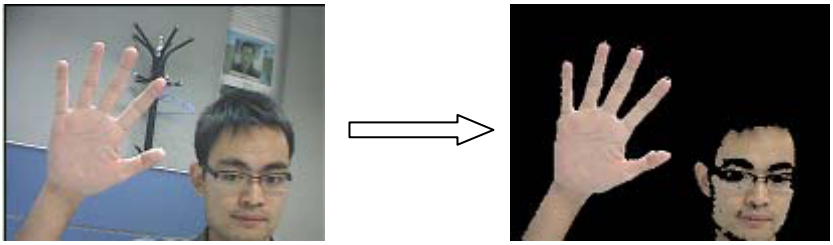


Fig. 1. Skin-color extracting using a single 2D Gaussian model. The left is a common image (RGB), the right is the corresponding skin-color image.

3 Representation of Hand Contour and Its Motion

In order to model the contour of a hand in an image, we use a graphic tool called B-spline curve. A B-spline curve is a parametric curve that allows representation of smooth, natural looking shape by specifying the basis functions and a small number of "control points". Suppose the coordinates of the control points are (x_1, y_1), (x_2, y_2) ... (x_n, y_n), then the B-spline curve can be expressed in matrix notation as follows.

$$(x(s), y(s))^T = B(s)\,(Q^x, Q^y)^T. \tag{1}$$

where s is the parameter, $B(s)$ is a 2×2n matrix called *metric matrix* whose entries are B-spline basis functions which are polynomials in s, and Q^x, Q^y are n×1 column vectors containing the x- and y-coordinates of the control points respectively. More details about B-spline parametric curves can be found in [12].

Similar to [8], we use a B-spline curve constructed from 50 control points as a deformable template for the hand contour. In the model, each finger and the two segments of the thumb are capable of rotating independently around their pivots respectively, and each finger except for the thumb is also capable of changing independently its length. The thumb is a bit different from the other fingers in that it consists of two segments which can rotate around their pivots but always keep a

constant length. This is because the thumb is assumed to flex only on the same plane as the palm. Since we assume that the user always keeps his palm approximately parallel to the image plane when moving in front of an ordinary camera, the hand contour model seems suitable for our tracking task. Besides the 50 control points, there are 6 pivots and a palm centre around which the whole hand rotates should be chosen. Figure 2 shows the deformable hand contour template.

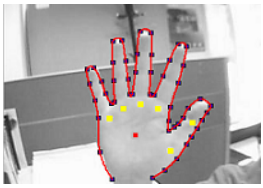


Fig. 2. Hand contour template showing its 50 control points, 6 pivots and 1 palm centre

Since we assume the user always keeps his palm approximately parallel to the image plane, the contour of his hand can only perform three kinds of movements. The first is rigid movement of the whole contour, including translation, rotation and scaling, the second is rotation of each finger's contour around their corresponding pivot, and the third is flexion/extension of the contour of each finger except for the thumb since we assume its two segments both keep a constant length. In conclusion, the movements of the hand contour can be represented by the following 14 dimensions state vector:

$$\chi = (x, y, \alpha, \lambda, \theta_L, l_L, \theta_R, l_R, \theta_M, l_M, \theta_I, l_I, \theta_{Th1}, \theta_{Th2}). \quad (2)$$

This state vector is partitioned into 7 parts as follows, the first part is the sub-vector (x, y, α, λ), which is a non-linear representation of a Euclidean similarity transform applied to the whole hand template. The second part is the sub-vector (θ_L, l_L) which represents the non-rigid movement of the little finger, θ_L means the little finger's angle with respect to the palm, and l_L means the little finger's length relative to its original length in the hand template. The following three parts are (θ_R, l_R), (θ_M, l_M), and (θ_I, l_I), they all have the same explanation as the sub-vector (θ_L, l_L), but for the ring finger, the middle finger and the index finger respectively. The sixth part is the parameter θ_{Th1}, which represents the angle of the first segment of the thumb with respect to the palm. And the last part is θ_{Th2}, which represents the angle of the second segment of the thumb with respect to the first segment of the thumb. It is noticeable that the finger and thumb angles are 0 when they are in the template position, and that the length of the fingers is relative to the template's finger length. So it's 1 when they have the same length as in the template. In our implementation, some constraints as follows are imposed on the finger lengths and finger angles, the minimum allowed length is 0.3 and the maximum is 1.2, and the finger angles must satisfy the following inequality: $-0.20 \le \theta_L \le 0.45$, $-0.25 \le \theta_R \le 0.25$, $-0.25 \le \theta_M \le 0.25$, $-0.45 \le \theta_I \le 0.20$, $-0.45 \le \theta_{Th1} \le 0.90$, $-0.10 \le \theta_{Th2} \le 0.25$.

4 Tracking the Contour of an Articulated Hand Throughout a Video Sequence

Two major techniques are used in our tracker, one is a kind of particle filter known as Condensation, which is introduced by Blake and Isard in [5]. The other is partitioned sampling, it is first introduced by MacCormick and Blake in [6] and then used by MacCormick and Isard [7] to implement a vision based articulated hand tracker.

4.1 The Condensation Algorithm Applied to Hand Contour Tracking

In section 3, we introduce a 14 dimensions non-linear state space which represents the movements of the hand contour. Each state vector in this state space determines uniquely a hand contour configuration deforming from the template. During the tracking process, for each frame, the task is to find the state vector in the 14 dimensions state space whose corresponding hand configuration fits best the real hand contour. The Condensation algorithm tells us how to find out such a state vector.

Suppose the state of the hand contour at time t is denoted as x_t and its history is $\chi_t = \{x_1, \ldots x_t\}$, similarly the observation at time t is denoted as z_t with history $Z_t = \{z_1, \ldots z_t\}$. The information for the location of the hand contour is expressed as a posterior conditional probability $p_t(x_t \mid Z_t)$, this is the probability of a hypothesized contour given the history of observations. The tracking objective is to find the state vector which maximize $p_t(x_t \mid Z_t)$. In general, it is difficult to calculate $p_t(x_t \mid Z_t)$ directly, especially when the dimension of the state space is high. For this reason the Bayes' theorem is applied to each time-step, obtaining the *posterior* $p_t(x_t \mid Z_t)$ based on all available information:

$$p_t(x_t \mid Z_t) = p_t(z_t \mid x_t)\, p_{t-1}(x_t \mid Z_{t-1}) \,/\, p_t(z_t). \tag{3}$$

where $p_{t-1}(x_t \mid Z_{t-1})$ is called the *prior*, and $p_t(z_t \mid x_t)$ is the *observation density*. As usual in filtering theory, a motion model between time-steps is adopted, this takes the form of a conditional probability density $p_t(x_t \mid x_{t-1})$ termed *dynamics*. Using the dynamics, equation (3) can be re-written as:

$$p_t(x_t \mid Z_t) = \frac{p_t(z_t \mid x_t) \int p_t(x_t \mid x_{t-1})\, p_{t-1}(x_{t-1} \mid Z_{t-1}) dx_{t-1}}{p_t(z_t)}. \tag{4}$$

Since $p_t(z_t)$ is generally a constant independent of x_t for a given image, it can be neglected so that only relative likelihoods need to be considered:

$$p_t(x_t \mid Z_t) \propto p_t(z_t \mid x_t) \int p_t(x_t \mid x_{t-1})\, p_{t-1}(x_{t-1} \mid Z_{t-1}) dx_{t-1}. \tag{5}$$

Formula (5) suggests that the conditional probability of the hand configuration at time t can be approximated as a sum of the previous conditional probabilities of the hand configuration multiplied by the dynamics, all weighted by the observation density.

In Condensation, an important notation called *weighted particle set* is used to represent a complicated probability distribution. A weighted particle set is a list of n

pairs (x_i, π_i) (i = 1, ... n) drawn from a probability density, where x_i belongs to a certain state space, and $\pi_i \in [0,1]$ is a weight on x_i proportional to the value of the density function at x_i, with $\Sigma^n_{i=1}\pi_i = 1$. Generally, any probability density can be represented by a particle set, no matter it is unimodal or multimodal.

In the context of hand contour tracking, since the posterior $p_t(x_t \mid Z_t)$ is usually too complex to be evaluated simply in closed form, we use a particle set to approximate it. Each particle drawn from the posterior comprises a state vector x_i which represents a hypothesized contour configuration, and a weight π_i which is the likelihood of the hypothesized contour representing the real target object. For the tracking task, we need to evolve in time the conditional density $p_t(x_t \mid Z_t)$, this is done in Condensation by propagating a particle set for the posterior from one time-step to the next. The propagation process in the Condensation algorithm consists of three operations called resampling, prediction and measurement, the following is a briefly description about them.

4.1.1 Resampling and Prediction

Suppose we already have a particle set for the posterior density at time t-1, the first operation need to be applied to this particle set is resampling. In Blake and Isard [5], this can be done efficiently with the use of cumulative probabilities. However, when the dimension of the state space is high their method does not work well, so we adopt another sampling strategy the same as described in Martin Tosas's thesis [8], details will be given in section 4.2.

After resampling, a certain motion model termed dynamics is applied to each resampled particle in order to generate some new states. This process is referred to as prediction. The mathematical fundamental behind the dynamics is the second-order auto-regressive processes (ARPs). As described in Blake and Isard [5], a second-order APRs model expresses the state x_t at time t as a linear combination of the previous two states adding on some Gaussian noise:

$$x_t = A_1 x_{t-1} + A_2 x_{t-1} + B w_t. \tag{6}$$

In equation (6), A_1, A_2 are fixed matrices which represent the deterministic components of the dynamics, B is another fixed matrix that represents the stochastic component of the dynamics, and w_t is a vector of independent random normal $N(0,1)$ variaets. In practice, we assume that the parameters of the 14 dimensions state vector (2) are independent mutually. So the matrices A_1, A_2 and B in (6) should be setted as diagonal matrices. In other words, the dynamical model (6) can be considered as 14 one-dimensional oscillators, one for each parameter of the articulated hand model, as described in [5]. A one-dimensional oscillator has the same form as (6) but with no matrix involved. For example, the oscillator which models the dynamics of the parameter θ_L can be written as:

$$(\theta_L)_t = a_1 \times (\theta_L)_{t-1} + a_2 \times (\theta_L)_{t-2} + b \times w_t \tag{7}$$

Blake [12] tells how to choose the coefficients a_1, a_2 and b:

$$\begin{aligned} &a_1 = 2\, exp(-\beta\tau)\, cos(2\pi f\tau),\ \ a_2 = -\, exp(-2\beta\tau), \\ &b = \rho\times[1-(a_2)^2-(a_1)^2-2\,a_2\,(a_1)^2/(1-a_2)]^{1/2} \end{aligned} \tag{8}$$

In equations (8), β is called damping constant, f called natural frequency and ρ called root-mean-square average displacement, the symbol τ is the time-step length in seconds. The oscillators for the other parameters have the same form as (7) but with different coefficients. In our implementation, all the coefficients are found empirically, though there are other approaches to obtain them as introduced in [12].

4.1.2 Measurement

After prediction, we obtain some new particles whose states are known, but do not have a weight yet. In order to calculate their weight, which is the likelihood of a particle (hypothesized contour) representing the real hand contour, the measurement operation is applied.

Given a new particle whose state vector is known, its weight can be calculated by the observation density $p_t(z_t \mid x_t)$. Though there are various forms of observation density to be chosen, the one introduced by Blake and Isard [5] are simple and effective so we adopt it in our implementation, but make some changes to it.

As do in Blake and Isard's approach, several line segments normal to the contour represented by a new particle at some well chosen measurement points are used to calculate a weight for the new particle. These line segments, termed measurement lines, are processed in order to find image features along them. Each of the found features has a contribution towards the weight of the measured particle. To quantify this contribution, we use a single Gaussian centred on the measurement point to model the features along each measurement line, as Blake [5] introduced. Sometimes, more than one detected features appear along a certain measurement line. For efficiency, only the one nearest from the measurement point is considered.

Currently, to measure a hypothesized contour, we only need to measure along each chosen normal line using a single Gaussian and multiply the results. By simplifying, the weight for a measured particle can be computed using the following formula:

$$\pi \propto p(z_t \mid x_t) \propto \exp\{-\sum_{i=1}^{M} \frac{1}{2\sigma^2} f(z_1(s_m) - r(s_m); \mu)\} \cdot \tag{9}$$

where π is the weight need to be evaluated, $f(v;u) = \min(v^2,u^2)$, $\mu = \sqrt{2\sigma}\log(1/\sqrt{2\pi}\,\alpha\sigma)$, $z_1(s)$ is the closest feature to the hypothesized contour along the measurement line s, and $r(s)$ is the position of the hypothesized contour along the measurement line s. To specify μ, three other parameters are needed: σ, the standard deviation of the single Gaussian mentioned above; $\alpha = q\lambda$ where q is the probability that the hand contour is not visible, and λ is the spatial density of the background clutter. More details about these parameters can be found in [12].

In Blake and Isard [5], all the measurement lines have the same length. This may suitable for their application, but for articulated hand tracking, it performs not very well, since the fingers of a hand are closed to each other, one may interfere with another, especially at the position closed to the root of a finger. To improve this situation, we adopt another strategy, namely, the length of the measurement lines for each finger are different, the closer the measurement point from the root of a finger, the shorter the measurement line at that point is. Experiments show that this strategy does reduce the interference between fingers. Figure 3 gives a comparison between two strategies.

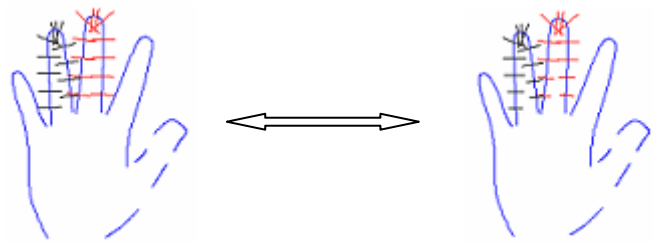


Fig. 3. Measurement lines: The left shows Blake's method, all measurement lines have the same length. The right shows our method, different measurement lines have different length.

4.2 Partitioned Sampling Used in Hand Contour Tracking

As proved in [7], more dimensions means more particles for Condensation, rendering the Condensation filter ineffective. So in our tracker, where the state space is too large to be dealt with a single Condensation filter, another technique called partitioned sampling is adopted to improve the efficiency.

In section 3, the state space is partitioned into 7 parts similar to [8].In this section, each of the 7 parts will have a corresponding sub-particle set, called *Palm*, *L*, *R*, *M*, *I*, T_1 and T_2 sub-particle set. Based on this decomposition, the partitioned sampling is performed hierarchically on the particle set to be evolved at each time-step, as follows:

1. For the *Palm* particle set do:
   1.1 Use each complete particle of the previous time-step to generate a number (250) of new particles for *Palm*, proportional to the weight of the used complete particle (Resample).
   1.2 Apply dynamics to each particle in *Palm* and then measure them.
   1.3 Select particles from *Palm* that constitute peaks of weight (top 10%).
2. For each of the finger particle sets, i.e. *L, R, M, I, and* T_1 do:
   2.1 Use each of the selected particles in *Palm* to generate a number (100) of new particles in the finger particle set, proportional to the weight of the used *Palm* particle (Resample).
   2.2 Apply dynamics to each newly generated finger particle and measure them.
   2.3 Group the newly predicted finger particle set based on which *Palm* particle they are generated from, and then select the particles with the highest weight from each group.
3. For the second thumb segment particle set, i.e. T_2 do:
   3.1 Use the select particles in T_1 to generate a number (100) of new particles in the T_2, proportional to the weight of the used particle in T_1 (Resample).
   3.2 Apply dynamics to each particle in T_2 and then measure them.
   3.3 Group the new T_2 particle set based on which T_1 particle they are generated from, and then select the particles with the highest weight in each group.
4. Form complete particles for the whole hand contour, which will be used in the next time-step, and choose the one with the highest weight as the tracking result.

A complete particle is a particle for the whole state vector. It contains 7 sub-particles, one for each of the 7 parts of the state vector, where the L, R, M, I and T_1 sub-particle are generated from the *Palm* sub-particle, and the T_2 sub-particle is generated from the T_1 particle.

In Martin Tosas [8], after step 2.2 and 3.2, a technique called particle interpolation is used so that it is convenient for forming complete particles in step 4. Instead of particle interpolation, we simply choose the best finger particles from each sub-particle set generated from the same "parent". Obviously, this is more effective, and experiments show that the result is satisfactory.

After step 4, we obtain several complete particles. To measure each of them as a whole, the following formula is used:

$$W=(2W_{\mathrm{Palm}}/8)\times(W_{\mathrm{L}}/8)\times(W_{\mathrm{R}}/8)\times(W_{\mathrm{M}}/8)\times(W_{\mathrm{I}}/8)\times(W_{\mathrm{T1}}/8\times)\times(W_{\mathrm{T2}}/8). \quad (10)$$

where W is the weight for a complete particle, and W_{Palm}, W_{L}, W_{R}, W_{M}, W_{I}, W_{T1}, W_{T2} are weights for the seven sub-particles respectively. Since the *Palm* sub-particle set is the first hierarchy to search, the generation of the rest sub-particle set is based on it, we assign a larger proportion to the *Palm* particle seems reasonable, experiments also indicate this is right.

Fig. 4. Some snapshots of our experiments

5 Experiments Results

Our tracker is able to track in real-time (30 frames per second) an articulated hand contour on a current common PC (**P**4 1.7G CPU and 512M memory) equipped with an ordinary camera (the image size is 320×240), both rigid motion and non-rigid motion can be tracked robustly, it performs well even in a clutter background. Figure 4 are some snapshots of our experiment.

6 Summarization

In this paper, a robust articulated hand contour tracker is introduced. The tracker works in real-time against a background without special requirements, it is able to

recover from miss-tracking automatically in some situations. The last three pictures in figure 5 indicate that it's not sensitive to interference, since it performs well even another hand put in front of the tracked hand. As the hand contour is modeled by B-spline curve, it is able to adapt most people's hand. In conclusion, the tracker performs well when tracking both rigid and non-rigid motion of anyone's hand contour.

For our tracker, the only used clue is skin-color. Other image features such as edges and frequencies are useful but neglected in our implementation. The configuration constraints between fingers are also useful information. In future work, these clues may be considered to improve the performance.

References

1. Rehg, J., Kanade, T.: Digiteyes: Vision-based hand tracking. Technical Report CMU-CS-93-220, Carnegie Mellon Univ. School of Comp. Sci (1993)
2. Kuch, J., Huang, T.: Vision-based hand modeling and tracking for virtual teleconferencing and telecollaboration. In: Proc. IEEE int. Conf. Computer Vision, pp. 666–671 (1995)
3. Stenger, B., Mendonca, P., Cipolla, R.: Model-Based 3D Tracking of an Articulated Hand. CVPR II, 310–315 (2001)
4. Stenger, B., Arasanathan, T., Torr, P., Cipolla, R.: Model-Based Hand Tracking Using a Hierarchical Bayesian Filter. PAMI 28(9), 1372–1384 (2006)
5. Isard, M., Blake, A.: Condensation–conditional density propagation for visual tracking. Int. J. Computer Vision 29, 5–28 (1998)
6. MacCormick, J., Blake, A.: A probability exclusion principle for tracking multiple objects. In: Proc. 7th International Conf. Computer Vision, pp. 572–578 (1999)
7. MacCormick, J., Isard, M.: Partitioned sampling, articulated objects, and interface-quality hand tracking. In: European Conf. Computer Vision (2000)
8. Tosas, M.: Visual Articulated Hand Tracking for Interactive Surfaces. PhD thesis, University of Nottingham (2006)
9. Nolker, C., Ritter, H.: GREFIT: Visual Recognition of Hand Postures. In: Proc. of the International Gesture Workshop, pp. 61–72 (1999)
10. Stefanov, N., Galata, A., Hubbold, R.: Real-time hand tracking with Variable-length Markov Models of behaviour. In: IEEE Int. Workshop on Vision for Human-Computer Interaction (V4HCI), in conjunction with CVPR (2005)
11. Kolsch, M., Turk, M.: Hand tracking with Flocks of Features. Computer Vision and Pattern Recognition, CVPR 2, 20–25 (2005)
12. Blake, A., Isard, M.: Active Contours. Springer, Heidelberg (1998)

Integrating Gesture Recognition in Airplane Seats for In-Flight Entertainment

Rick van de Westelaken, Jun Hu, Hao Liu, and Matthias Rauterberg

Technische Universiteit Eindhoven, Den Dolech 2
5600MB Eindhoven, The Netherlands
H.F.M.v.d.Westelaken@student.tue.nl,
{J.Hu,hao.liu,G.W.M.Rauterberg}@tue.nl

Abstract. In order to reduce both the psychological and physical stress in air travel, sensors are integrated in airplane seats to detect the gestures as input for in-flight entertainment systems. The content provided by the entertainment systems helps to reduce the psychological stress, and the gesture recognition is used as input for the content and hence stimulates people to move, which as a result would reduce the physical stress as well.

Keywords: In-flight entertainment, gesture recognition, air travel, air travel thrombosis.

1 Introduction

Today a large number of people use air travel as a way of transportation. This number is increasing every year, as well as the number of long-haul flights. At the same time the flight duration is increased because of better fuel efficient airplanes, which makes intermediate landings unnecessary. The air travel market is highly competitive and therefore airlines try to maximize the number of seats [1]. Often this results in a very limited amount of seating space for passengers, especially in economy class [2]. In this context the EU project "SEAT" (Smart tEchnologies for stress free Air Travel) was set-up [3]. This project, sponsored by a European commission focuses on improvements and solutions to increase comfort in air travel. The partners in this project are Imperial College London, Acusttel, Aitex, Antecuir, Czech Technical University, Wearable Computing Lab ETH Zürich, Inescop, Queen Mary University of London, Starlab, Eindhoven University of Technology, Thales and DHS. The work described here is part of the SEAT project and is executed in Eindhoven University of Technology.

The combination of long flight duration, limited space and an unusual cabin environment in terms of air pressure, humidity and continues noise causes physical and psychological discomfort for a large group of passengers as well as the crew [4]. Most of the airlines offer in-flight entertainment to their passengers. This provides mental distraction and might lead to reduction of psychological stress. To reduce physical stress some airlines recommend passengers in-flight exercises[2, 5]. The ways these exercises are presented vary from paper versions to complete instruction

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 353–360, 2008.

videos and crew demonstrations. The main goal of these exercises is to stimulate the blood flow and prevent health related issues like deep vein thrombosis (DVT), stiffness and fatigue. The problem with these recommendations is that it depends on the passenger whether the exercises are executed. Another way to prevent DVT is to wear compression stocking to prevent formation of DVT [6]. Ordinary this is only used by passengers with increased risk of DVT, which means that only a relative small group of people use this means.

Physical stress can cause a number of health related problems, as shown by different institutes like the World Health Organization [7] and the British Medical Association [8]. Therefore this work focuses on how passengers can be motivated to move during long-haul flights in order to reduce physical stress. The solution discussed here is an airplane chair equipped with sensors to detect the body movements and gestures of passengers. The detected gestures could than be used as input for interactive applications in in-flight entertainment systems.

The idea of controlling interactive content by means of body movement is not new. In contrary it is very popular at the moment. Examples of movement controlled game platforms are the Nintendo Wii and Sony Eye-Toy. Although there is a big difference between these platforms, these platforms assume that the user has enough space in which they can move. This assumption does not hold in an economy class flight environment. There the amount of space is very limited, which means that new solutions had to be explored.

2 Problem and Context

Prolonged immobility by sitting in a chair during long haul flights can lead to pooling of blood in the legs. It is known that immobility cause formation of blood clots in the body and especially in deep veins (Deep Vein Thrombosis, or DVT in short). *"Larger clots may cause symptoms such as swelling of the leg, tenderness, soreness and pain. Occasionally a piece of the clot may break off and travel with the bloodstream to become lodged in the lungs. This is known as pulmonary embolism and may cause chest pain, shortness of breath and, in severe cases, sudden death. This can occur many hours or even days after the formation of the clot."* [7]

A well known way to reduce the risk on DVT is by wearing compression stockings. Compression on the leg surfaces forces blood to flow from the small surface vessels into the larger, deep venous system. Besides it also prevents back-flow of blood and the formation of clots. Generally only passengers with increased chance on DVT use these stockings. Although these stockings reduce risk on DVT it does not deal with physical discomfort in general.

In order to reduce physical discomfort contraction of muscles is very important. Muscle activity helps to keep the blood flowing through the veins, particularly in the deep veins. A way of generating muscle activity is by moving around in the cabin. However people don't want to cause inconvenience to neighbor passengers while passing or are too tired. Also potential health benefits should be balanced against the risk of allowing a lot of passengers walking around in the plane with possible chance on unexpected turbulence. For this reason it is recommended to stimulate muscle activity while sitting in a chair. Many Airlines already provide a number of exercises

to their passengers which can be performed during the flight. Research proved that the most effective exercises to stimulate circulation and thereby reduce discomfort, fatigue, stiffness and DVT. Research has been done to evaluate recommended lower leg exercises by airlines to investigate which exercises induce optimum calf muscle pump activity[5], and these exercise provide important information as starting point for a solution (Fig. 1).

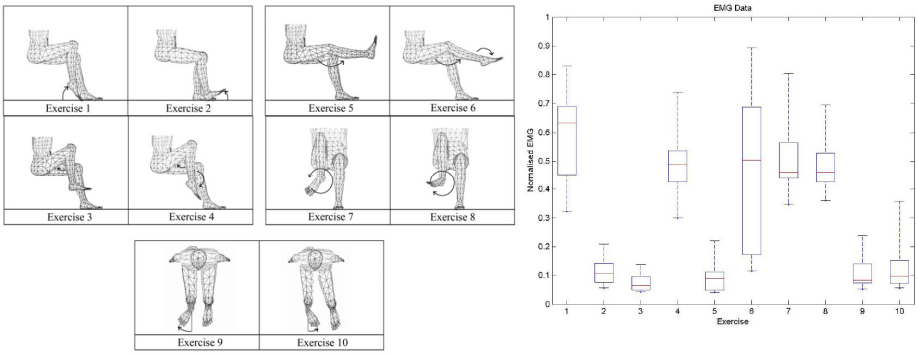


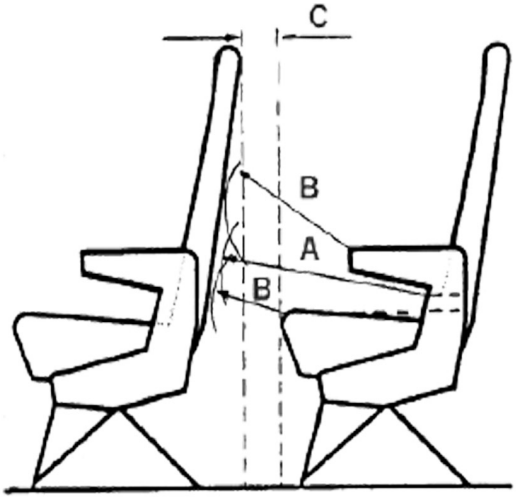
Fig. 1. Recommended exercises and produced EMG in calf muscle [5]

In an economy class cabin environment of an airplane, the dimensions of the chairs and space between the chairs are an important factor towards a solution. There are no global regulations concerning seat spacing in airplanes. The "Civil Aviation Authority" (CAA) based in the United Kingdom formulated regulations concerning seat spacing for planes registered in the UK. These Regulations entitled "Airworthiness Notice No. 64 (AN64)" [6] dates from 1989 and is not revised since. Currently these regulations are still used as leading regulations for seat spacing within JAA countries (Countries connected to the Joint Aviation Authorities, which include a large number of European countries). The seat requirements are split into three leading dimensions, shown in Fig. 2.

These dimensions introduce a real challenge because when the AN64 regulations were made it satisfied 95 percentile for dimension "***A***". Nowadays the same regulation only satisfies 77 percentile of European passengers as result of increased average length of Europeans.

The AN64 guidelines do not describe a minimum width of the chairs. An investigation to chair widths among different airlines in economy class shows a minimum of 420 mm and a maximum width of 480 mm. Out of the seat measures in economy class can be concluded that this is a big constraint for in-chair exercises.

Today in-flight entertainment is offered to reduce psychological stress. This project on the other hand focuses on reducing physical stress. By reduction of physical stress the chance on health related issues like DVT, stiffness and fatigue are reduced [9]. In order to achieve this, passengers are required to exercise. These exercises can be described as a set of simple movements. However the real challenge is to stimulate passengers to make these movements. To reach this goal the in-flight entertainment system could by used or it could simply be an expansion of the in-flight entertainment.



Dimension	Description	Minimum
A	The minimum distance between the back support cushion of a seat and the back of the seat or other fixed structure in front	26 inches (660 mm)
B	The minimum distance between a seat and the seat or other fixed structure in front.	7 inches (178 mm)
C	The minimum vertically projected distance between seat rows or between a seat and any fixed structure forward of the seat.	3 inches (76 mm)

Fig. 2. Seat dimensions according to AN64 [10]

3 Concepts

With the problem and the context as starting points, an idea generation session was held with four industrial designers. The most promising ideas generated during this session are shortly described below. A remarkable detail is that all ideas except one are based on games because games were considered to be typical interactive content that motivates people to take active roles [11].

1. Quiz game; Answering questions happens by body movement.
2. Drive game; with physical steering wheel and pedals.
3. Agility game; balancing digital objects by physical movement.
4. Copy game; try to copy movements of others as well as possible.
5. Movement of passengers will be triggered by changes somewhere in the cabin.

A benchmark based on a Pugh matrix (also called a decision matrix) has done to find the most promising idea. A Pugh matrix is a way to assess ideas based on pre selected criteria and matching weight factors. The result of this matrix should show the most promising idea(s) according to the set criteria. The used assessment criteria from most important to less important are:

1. Usability of basic principle; can the basic principle behind the idea be used for a wide variety of games / exercises or is it only applicable for a small range of games and exercises.

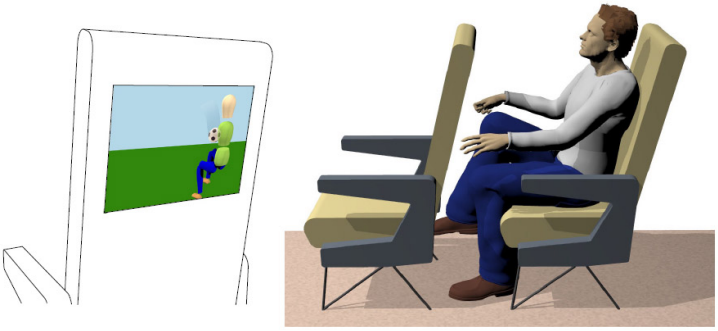


Fig. 3. Control in-flight entertainment by body movement

2. Feasibility within space; to what extend might the limited amount of space in the plane' chair lead to problems in executing the exercises.
3. Movement intensity; to what extend do passengers have to move concerning the intensity and repetitions.
4. Disturbance of other passengers; Does the idea lead to disturbance or annoyance of other passenger.

The benchmark showed an almost equal score for a few ideas. The similarity among these ideas was that it contained digital content that is controlled by physical movement. The difference on the other hand is the digital content used for every idea. This means that a generic interaction platform could be a solid basis for a width range of applications. The differences would be made by different content. With this assumption the project continued with finding solutions on how to translate physical movements or gestures to digital information that could be used as input for digital content.

Solutions to measure movements or gestures can be found in two directions: with attached sensors or detached sensors.

Attached sensors

A set of first solutions shown in figure 3 are based on attached sensors (sensors that need to be attached to the passenger). The attached sensors are therefore in most cases loose objects. This introduces some feasibility problems, which can be summarized by three keywords: Safety, effort and dependency.

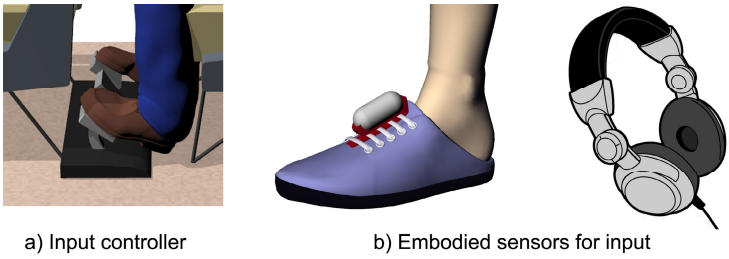


Fig. 4. Suggested detached sensors

Safety: Whether this are written rules or not, it is logical that loose objects will cause serious safety issues in certain circumstance. These circumstances could be turbulence or passengers who are diminished responsible that use it for other purposes.

Effort: In case that a technological solution is found in a loose (embodied) object the chance on success smaller. The user has to put in extra effort in order to use the system. For some users this threshold is too high for tem to use the system.

Dependency: Regular objects can be used to obtain data from the user. For example the standard headphones could be equipped with sensors to measure a passenger's head and neck movement. The problem with this is that a percentage of the passengers won't use this headphone or will bring their own. This results in uncertainty about usage of the objects, which could lead to failure of the system.

The usage of separate attached objects is not desired and therefore ideally the solution should be fully integrated in the airplane.

Detached sensors

A number of technical solutions with detached sensors were considered. These solutions are described below.

Video-based gesture recognition: This way of gesture recognition is not ideal in the context of an economy class airplane. The biggest problem lies in the fact that the amount of free space is very limited. When the camera(s) is placed too close to the passenger there is the problem of not being able to cover enough of the passenger's body.

Sensors integrated in the floor: With sensors integrated in the plane's floor it is possible to measure the angle and distance of a passenger's feet in relation to the floor. With this data the user's lower body position can be calculated. This could be quite accurate because human anatomy and the usable movement space are known. However it would become integrated in the plain which makes adaptations or repairs very difficult and costly.

Sensors integrated in the chair: By a grid of pressure sensors integrated in an airplane chair it is possible to measure a person's weight distribution. When there are changes in this weight distribution it is possible to derive a person's position or gesture out of this data.

The last suggestion seems to be the most promising solution. However the question is how many sensors are needed to obtain a certain accuracy level and how many gestures can be recognized. Since the main goal is to stimulate people to move, motivating people to move is the more important than knowing exactly what movements are made. The system might work while it can only recognize a few different gestures. Therefore a system with a relative low accuracy might be sufficient. However this assumption needs to be further validated.

4 Solution and Prototyping

The suggested solution is an airplane chair equipped with a grid of pressure sensors to detect a passenger's body weight distribution. The seating surface is the most

important part because there is always contact between this surface and the passengers. At the same time the weight changes in this surface are relative large compared to weight changes in the back area. This does not mean that sensors in the back surface are useless. This could be a good enrichment of the other data obtained from the seating surface. The obtained sensor values are processed and by means of pattern recognition the corresponding gestures are derived. When the gesture is determined this information can be used as input for the in-flight entertainment systems.

To validate the concept and to prove its effectiveness a prototype was build. Starting point for this prototype is a chair made out of MDF with similar dimensions of an average airplane chair. The seating surface of this chair is eventually equipped with a grid of 28 FSR's (Force Sensing Resistor). In between the FSR and the user a four millimeter thick rubber plate is placed. On top of this rubber a round plate with a diameter of 40 millimeter is placed. This top plate is used to increase the surface that press on the FSR. The rubber plate is used to equally divide the weight on the FSR. The grid of sensors is connected with a 4 to 16 line decoder to a microcontroller. The data is send to a computer by means of a serial data connection. It takes about 50 ms. to read all the sensor values and to send them to a computer. The sensor values are fed into a neural network and the output of this neural network is the corresponding gesture.

For implementing the neural network a software package "JOONE" was used. JOONE (Java Object Oriented Neural Engine) is a Java based application to simulate neural networks and to train them [12, 13]. The neural network used for the prototype is a three-layer sigmoid network that consists out of 28 input neurons. In-between the input and output layer there is one hidden layer consisting out of 30 neurons. The output layer, which has nine neurons represent the different gestures.

The prototype is able to recognize nine gestures. A first test application had been created to test the principle of a gesture controlled game, where users have to balance a ball on a tray by lifting their legs was created. This game showed that a capability of only detecting three gestures (lifted right leg, lifted left leg, both legs down) already worked quite well.

Fig. 5. First prototype

5 Conclusion and Future Research

To reduce both the psychological and physical stress in air travel, the work presented here suggests a solution of integrating sensors in airplane seats to detect the gestures as input for in-flight entertainment systems. The content provided by the entertainment systems would help to reduce the psychological stress, and the gesture recognition is used as input for the interaction and hence stimulates people to move which, in turn reduces the physical stress as well.

The future research is needed to validate the assumption that this concept can really stimulate people to move. If so the next research question would consider the effectiveness of the concept. Do passengers move enough to reduce physical discomfort compared to current situations? And is the chance on health related issues like thrombosis really smaller by using this chair? A series of user studies has been planned and will be carried out to stress these questions.

References

1. Quigley, C., et al.: Anthropometric Study to Update Minimum Aircraft Seating Standards, EC1270, prepared for Joint Aviation Authorities. ICE Ergonomics Ltd (2001)
2. Hinninghofen, H., Enck, P.: Passenger well-being in airplanes. Auton Neurosci. 129(1-2), 80–85 (2006)
3. SEAT Project Consortium, SEAT Project (2006), `http://www.seat-project.org`
4. Hickman, B.J., Mehrer, R.: Stress and the effects of air transport on flight crews. Air Medical Journal 20(6), 2–56 (2001)
5. O'Donovan, K.J., et al.: An investigation of recommended lower leg exercises for induced calf muscle activity. In: Proceedings of the 24th IASTED international conference on Biomedical engineering, 2006 of Conference, pp. 214–219. ACTA Press (2006)
6. Ball, K.: Deep vein thrombosis and airline travel–the deadly duo. AORN Journal 77(2), 346–354 (2003)
7. World Health Organisation, Chapter 2: Travel by Air: Health Considerations, International travel and health: Situation as on 1 January (2005), `http://www.who.int/ith`
8. Dowdall, N., Evans, T.: The impact of flying on passenger health: a guide for healthcare professionals, BMA policy report, Board of Science and Education, British Medical Association (2004)
9. Alonso, M.B.: Affective Tangible Interaction; Towards Reducing Stress. In: Proc. HCI Close 2U - 9th Sigchi.nl conference (2005)
10. Civil Aviation Authority, CAP 747: Mandatory Requirements for Airworthiness (2006), `http://www.caa.co.uk/docs/33/CAP747.PDF`
11. Breuer, H.: Interaction Design for Flight Entertainment, bovacon (2006)
12. Jeff, T.H.: Introduction to Neural Networks with Java. Heaton Research, Inc. (2005)
13. Marrone, P.: The Joone Complete Guide (2007), `http://www.joone.org`

Designing Engaging Interaction with Contextual Patterns for an Educational Game

Chien-Sing Lee

Faculty of Information Technology, Multimedia University, Cyberjaya 63100 Selangor, Malaysia
cslee@mmu.edu.my

Abstract. This paper aims to address two problems. The first problem is how to develop engaging (deep and meaningful) pedagogical patterns and still ensure there is sufficient incremental cognitive complexity. The second problem deals with how to connect pedagogical patterns to HCI and software engineering to form a systemic interaction design framework. An educational game is used as an example. Significance of the study lies in the development of a means to create interaction designs around design-for-engagement requirements and the flexible scaling and synergy of different frames of reference (pedagogy-application domain-HCI- software engineering) through instantiations from the synergised patterns; reducing error and cost and encouraging new experimentations with transfer of engaging learning.

Keywords: Deep and Meaningful Interaction, Educational Game, Contextual Pattern-based Interaction Design, Pedagogical Structure, Scalability

1 Introduction

Instructional design strategies revolving around the iterative Analyze, Design, Develop, Implement and Evaluate (ADDIE) model have been regarded as the conventional framework that guide the design of systematic learning interaction. At a higher level of design, interaction tasks are mapped to learning outcomes. Bloom's revised taxonomy (remembering, understanding, application, analysis, evaluation, creation)[1] and Kirkpatrick's four levels of evaluation (reactions, learning, transfer in behavior and effects on business results) [2] are often cited as references that help this mapping. The former is for teaching and learning and the latter for training purposes. However, deep and meaningful interaction design often remains elusive. To address this problem, we will first look at what constitutes deep and meaningful interaction.

Earlier work has classified interactions based on the learning outcome when students interact with specific media such as hotspots or animations [3, 4]. However, [5] believes that quality of interaction should consist of a greater number of interactions involving deep and meaningful cognitive processing such as problem solving, decision-making and evaluation compared to those of lower levels of cognitive processing. Another perspective to interaction design measures its success by the degree that it enhances usability and consequently, user experience. Usability is assessed in terms

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 361–370, 2008.

of the degree of effectiveness, efficiency, safety, utility, ease in learning and ease in remembering. If these core factors are met, then user experience is more likely to result in interactions that are fun, helpful, satisfying, motivating and encourages creative expressions [6]. Hence, deep and meaningful interaction extends from human-computer interaction issues such as how to create user-friendly means for keying in input, navigation or search. Deep and meaningful interaction involves higher cognitive processes mediated through high degree of usability and satisfying user experiences that motivate and encourage creative expressions.

1.1 Problem Statements

Two problems are addressed. First, successful practices in designing for engaging (deep and meaningful) interaction need to be shared, adopted or adapted for further innovations to occur. As such, it is essential to identify a means to guide development of interaction design for engaging learning.

Second, research on pedagogical patterns and interaction design patterns have often been carried out separately from each other as the focus of the former is often on identifying recurring problems and successful solutions for teaching and learning, whereas the latter is concerned with successful solutions to recurring usability and user experience problems with technology as mediator in human-computer interactions. Factored into any systems design, the software engineering aspect need to be added to pedagogical and human-computer interaction considerations. Hence, in order to design deep and meaningful interaction from a systems perspective, there is a need to synergize pedagogy, human-computer interaction and software engineering to form a common frame of reference so that the interdisciplinary patterns obtained through the tri-partite synergy can become the basis for forming pedagogically oriented interdisciplinary pattern languages.

1.2 Significance of the Study

Incorporating design-for-engagement guidelines into the synergized framework provides a basis for assessing and validating designs prior to development. In addition, these design-for-engagement guidelines can be used to allow flexibility in the adaptation of successful patterns for traditional learning environments to less structured interactive media such as games-based learning.

Furthermore, a synergised pedagogical-application domain-HCI-software engineering framework will enable easier factorization of the analysis, design, development and evaluation of and communication between one discipline and other disciplines. Having a common frame of reference at the onset ensures sufficient rigor and co-designer participation by all stakeholders in the analysis and design phases prior to development.

The organization of this paper is as follows: Related work on the student-to-content interaction design matrix, pattern approach to interaction design and pedagogical patterns are first reviewed. Subsequently, a synergised PAHS (pedagogical- application domain-HCI-software engineering) framework for designing solutions to the three problems raised is presented interspersed with an educational game example.

2 Related Work

2.1 Bloom-Content Interaction Mapping

Quality interaction design is reflected in the level of cognitive engagement as concluded above. Although Bloom's revised taxonomy provides the basis for planning and assessing incremental cognitive complexity in instructional design, there is a need to map learning activities to the type of content interaction so that there is an explicit correlation between learning activities and the type of HCI content interaction. [7] classifies content interaction types into ten categories: enriching, supportive, conveyance, constructive, triggering, exploration, integration, resolution, reflective inquiry and metacognitive. These interactions are applicable across Bloom's revised taxonomy. However, in Table 1 below, the author maps the types of content interaction by [5] to the cognitive process that will most benefit from its use. The author also includes evaluation as an activity (in italics) for integration interaction. Hence the difference between integration and resolution interactions is in the development of novel ideas in the latter.

Table 1. Mapping between Bloom's revised taxonomy and categories of interaction

Bloom	Content interaction type
Remembering	*Triggering* interactions - create interest in learning
Understanding	*Conveyance* interactions - demonstrate *Exploration* interactions - encourage learners to determine their own learning path and search deeper into the area of interest *Enriching* interactions - enable access to information such as hyperlinking to additional resources. *Supportive* interactions - aid experimentation such as zooming in and searching and querying.
Application	*Conveyance* interactions – allow students to apply their knowledge, for instance through simulations or games
Analysis	*Constructive* interactions - encourage active participation in organizing and mapping knowledge to reflect the learner's understanding such as by drawing cognitive maps. *Integration* interactions - establish relationships between ideas and develop solutions
Evaluation	*Integration* interactions - establish relationships between ideas and develop solutions *and evaluate solutions*
Creation	*Resolution* interactions - develop new ideas and evaluate these solutions

Similarly, Merrill recommends guidelines [8] in his First Principles of Instruction with engagement as the core for designing instruction. These guidelines are formed from the commonalities among many instructional models:

- Learning is facilitated when learners are engaged in solving real-world problems (learning by doing)
  - o Learners are to be involved in problem identification and not merely problem-solving

  - o Learners need to be shown the task that they are going to solve
  - o Learners need to explicitly identify differences from one stage of learning to the other
- Learning is facilitated when existing knowledge is *activated* as a foundation for new knowledge (learning by doing)
  - o Learners are asked to recall, relate, describe or apply knowledge from past experience
  - o Learners are provided with relevant experience fundamental to the next task
  - o Learners are given the opportunity to demonstrate their grasp of knowledge
- Learning is facilitated when new knowledge is *demonstrated* to the learner (careful sequencing of learning activities)
  - o Demonstration is consistent with the learning goal e.g. demonstrations of procedures, visualizations of processes and modelling of behaviour
  - o Learners are given suitable guidance e.g. provision of relevant information, multiple forms of representations (text, graphics, videos etc.), contrasts between demonstrations
- Learning is facilitated when new knowledge is *applied* by the learner (reuse)
  - o Appropriate feedback and coaching should be provided inclusive of identification and correction of errors
  - o The problems to be solved should be varied but incremental in complexity.
- Learning is facilitated when new knowledge is *integrated* into the learner's world (reuse and sharing)
  - o Learners should be given the opportunity to demonstrate their new knowledge
  - o Learners are given opportunities to reflect, discuss and defend their opinion
  - o Learners are given opportunities to create and explore new ways to use their new knowledge

The author takes engagement in Merrill's First Principles as the core that directs other principles. As such, in Table 2, engagement is presented in the first column, mapped to corresponding principles, student-content interaction types and Bloom's taxonomy.

2.2 Pattern Applications to Interaction Design

The process of interaction design involves 4 activities: identifying needs and establishing requirements, developing alternative designs that meet those requirements, building nteractive versions of the designs so that they can becommunicated and assessed and evaluating what is being built throughout the process[6]. Patterns encapsulate the name, ranking (stretching from tried-and-tested to new), illustrations, the problem(s) addressed, forces (gap in design aspects that need to be improved), examples, solution(s), and diagram (summarized version of the illustration). Connecting related patterns form a pattern language. Some examples of popular patterns are listed in [9]'s design patterns.

From the perspective of usability engineering, [10] integrates application domain pattern languages, human-computer interaction pattern languages and software engineering pattern languages. The application domain pattern language deals with

Table 2. Mapping between First Principles, student-content interaction types and Bloom

First Principles (core)	First Principles	Interaction type	Bloom
Engaged – learners need to be involved in problem identification and not merely problem-solving	*Activated* - learners are asked to recall, relate, describe or apply knowledge from past experience - learners are provided with relevant experience undamental to the next task	*Triggering interactions* - create interest in earning	Remembering/ Recalling
Engaged – learners need to be shown the task that they are going to solve	*Demonstrated* - demonstration e.g. procedures, isualizations of processes and modeling of behavior - learners are given uitable guidance e.g. provision of relevant information, multiple forms of representations, contrasts between demonstrations	*Conveyance* interactions - demonstrate *Exploration interactions* –encourage earners to determine their own learning path and search deeper into the area of interest *Enriching interactions* - enable access to information such as hyperlinking to additional resources. *Supportive interactions* -aid experimentation such as zoom, search and query	Understanding
Engaged –learners need to explicitly identif differences from one stage to another	*Applied* - Feedback and coaching - Incremental complexity - Varied	*Conveyance interactions* - allow students to apply their knowledge, for instance through simulations or games	Application
	Activated - learners are given the opportunity to demonstrate their grasp of knowledge	*Constructive interactions* - encourage active participation in organizing and mapping knowledge to reflect the learner's understanding such as by drawing cognitive maps.	Analysis
	Integrated - learners demonstrate their new knowledge, reflect, discuss and defend	*Integration interactions* -establish relationships between ideas and develop solutions	Evaluation
	Integrated - learners create and explore new ways to use their knowledge	*Resolution interactions* - develop new ideas and evaluate these solutions	Creation

large-scale to small-scale concepts whereas the human-computer pattern language defines the tasks, dialogues and interaction objects; and the software engineering pattern language specifies the architecture, design and implementation guidelines. Both application and human-computer interaction pattern language designers communicate to contextualize the design requirements and process within a project environment and consequently, develop a suitable user interface. On the other hand, the human-computer interaction pattern language designers and the software engineering pattern language designers interact to produce the backend software design.

This tri-disciplinary integration fits well with the processes involved in interaction design mentioned above. However, [11] recommends that more user input should be factored in at the user interface design phase, shifting greater attention to the interaction aspect than the conventional emphasis on software design itself. Hence, he has applied Nielsen's [10] usability engineering model as the framework by which the three pattern languages above can be integrated and still fit the interaction design processes. The greater emphasis on HCI aspects is reflected in the usability-engineering model phases below. The pattern language applied is indicated in parenthesis:

1. Identifying whether the activities in a particular application domain can be captured as patterns. (application domain languages)
2. Comparing existing solutions for competitive analysis and transferring HCI pattern languages from successful competing products to the design context being analyzed. (HCI pattern languages)
3. Prioritizing usability measures such as learnability, efficiency of use, memorability and low error rate to determine necessary tradeoffs in model design. (HCI pattern languages)
4. Applying high-level HCI patterns as guidelines for designing initial prototypes. (HCI pattern languages)
5. Inviting application domain experts to provide feedback. (application domain languages)
6. Using low-level HCI patterns to provide the consistency in documentation, help systems and tutorials for the current product and existing products within the company. At this point, HCI patterns can serve to refine existing style guides and guidelines. (HCI pattern languages)
7. Developing prototypes guided by software patterns. (Software engineering pattern languages)
8. Using application domain patterns for constructing realistic application scenarios for empirical testing. (application domain pattern language)
9. Identifying design alternatives based on application domain, HCI and software patterns and refining the prototype iteratively. (application domain + HCI pattern language + software engineering pattern language)
10. Obtaining feedback from actual field use and applying application domain pattern languages to facilitate discussions between the user interface designers and users and HCI patterns to guide designers to alternative solutions. (application domain + HCI pattern language + software engineering pattern language)

Since patterns and pattern languages themselves are constantly evolving and improving, patterns provide the structure and the ontological basis for the design of interactive systems. However, architecturally, inclusion of other patterns such as

pedagogical patterns should be seamless, similar to inclusion of a component building block to software architecture. The following subsection introduces pedagogical patterns commonly accepted in educational circles.

2.3 Pedagogical Patterns

The Pedagogical Patterns project [12] was formulated during the OOPSLA'95 conference. Aimed at aiding novice instructors, who have good knowledge of their subject but are not necessarily experienced in teaching the subject, the Pedagogical Patterns provide general guidelines from successful teaching and learning strategies from which instructors can adapt based on their own creativity to suit their students' needs. The pedagogical patterns however, are not presented in the same format as design patterns.

In the section below, the application domain-HCI-software engineering framework is extended to include pedagogy.

3 Synergised PAHS Framework

The Pedagogical-Application domain-HCI-Software engineering (PAHS) framework builds on Borcher's [10] interrelated pattern languages framework and incorporates periodical user feedback and pedagogical pattern language within the interaction design [6] – usability-engineering [11] context. The focus of this paper is on the development and integration of user feedback and pedagogical patterns into PAHS.

3.1 PAHS Pedagogical Patterns

PAHS regards instructional design phases as seamlessly iterating between one and the other with no strict application in the choice and order of pedagogical techniques in order not to constrain creativity. What is essential is the observance of the need to analyze, design, develop and evaluate iteratively with as much user feedback as possible at each phase. In PAHS, pedagogical patterns are derived from a cross-section of useful guidelines such as those illustrated in Table 2 above. It is assumed that these patterns as well as the application domain, HCI and software engineering patterns can be retrieved from a database through a graphical user interface. An example of PAHS pedagogical patterns for an educational game in the domain of water pollution is shown in Table 3 below. In cases where elements in First Principles and tactics are similar, they are merged.

3.2 Pedagogical Patterns Integrated with the Usability Engineering Framework

PAHS includes user feedback and pedagogical patterns into Borcher's [10] integrated pattern languages framework within the user-centered interaction design [7] context. As a recap, interaction design requires first identifying needs and establishing requirements, second developing alternative designs that meet those requirements, third building interactive versions of the designs so that they can be communicated and assessed and fourth evaluating what is being built throughout the process.

Table 3. Example of PAHS pedagogical patterns for an educational game

First Principles	Content-interaction types	Tactics	Bloom's
Learners are to be involved in problem identification and not merely problem-solving	*Triggering* interactions - create interest in learning	Learners choose information which helps them to identify the problem(s), link and annotate those words to useful vocabulary from past experience and group them and click on one of the groups to start the game -feedback: After grouping, the user will be transported to a scenario which best fits the words	Remembering / Recall
Learners need to be shown the task that they are going to solve	*Conveyance interactions* -demonstrate for example, through simulations	Animate what happens to sea-life when there is an oil spill in the ocean Feedback: Ask students to use the tools available to clean up the oil spill	Understanding
Learners need to explicitly identify differences from one stage of learning to the other	*Exploration interactions* -encourage learners to determine their own learning path and search deeper into the area of interest	Learners decide on alternatives to clean up the oil spill and to reduce death of sealife.	Application
Learners need to explicitly identify differences from one stage of learning to the other	*Supportive* nteractions -aid experimentation	Enable students to experiment with the decision made through role-playing and adding different obstacles and rewards along the way.	Application
	Enriching interactions - enable access to extra information		
	Constructive interactions -encourage active participation in organizing and mapping knowledge	Learners use tools in the game to analyze the cause-effects of their decisions and reflect their analysis in a mind map	Analysis
	Integration interactions - establish relationships between ideas and develop solutions	Learners conclude and defend which line of action would provide the best results	Evaluation
	Resolution interactions - develop new ideas and evaluate these solutions	Learners develop new ways to solve the problem by creating their own path with the tools in the game	Creation

Supposing that the needs of the students is to learn the biological effects on sea life due to oil spills in the sea and to identify alternative ways to clean up the oil spills within the shortest time and least cost and least loss of sea lives, the following tasks are carried out instantiated from Nielsen's usability engineering framework:

1. Obtain user requirements (user feedback)
2. Identify activities in the application domain (application domain languages).
3. Validate these activities for engagement by mapping these activities to Merrill's First Principles, the student-content interaction types and Bloom as shown in Table 2 (pedagogical pattern language).
4. Obtain user feedback and refine activities. (user feedback and iterative refinement of activities/pedagogical pattern language)
5. Retrieve and compare existing solutions retrieved from the database based on the mastery level of the individual student. (application domain language)
6. Retrieve relevant and competing HCI pattern languages corresponding to the student-content interaction types above (HCI and application domain pattern languages).
7. Rank usability measures for example learnability, efficiency of use, memorability and low error rate to decide on necessary tradeoffs in model design. (HCI pattern languages)
8. Apply high-level HCI patterns to guide the design of initial prototypes. (HCI pattern languages).
9. Invite application domain experts and users to comment and revise activities and high-level HCI patterns as suggested. (user feedback, iterative refinement of activities, application domain, high-level HCI pattern languages)
10. Use low-level HCI patterns to provide the consistency in documentation, help systems and tutorials for the current product and existing products within the company. At this point, HCI patterns can serve to refine existing style guides as well as guidelines. (HCI pattern languages)
11. Obtain user feedback and refine low-level HCI patterns (user feedback and iterative refinement of activities/pedagogical patterns, application domain, high-level and low-level HCI pattern languages)
12. Develop prototypes using software architectural and/or design patterns. (Software engineering pattern languages)
13. Use application domain patterns to develop real-world scenarios for field testing. (application domain pattern language)
14. Identify design alternatives based on application domain, HCI and software patterns and refine the prototype iteratively. (application domain, HCI and software engineering pattern language)
15. Obtain feedback from experimental field use and apply application domain pattern languages to facilitate discussions between the user interface designers and users and use HCI patterns to guide designers to alternative solutions. (user feedback, iterative refinement of activities/pedagogical patterns, application domain, high-low-level HCI, software engineering pattern languages)

4 Conclusion

Effective learning is the ultimate goal of any instructor. One of the main essences of effective learning is the development of engaging learning. As such, this paper is concerned with the design of pedagogically oriented engaging interaction design. The author has presented the development of a user-centered synergised and pedagogically

oriented pattern language framework, the Pedagogy-Application domain-HCI-Software engineering framework or PAHS and applied the PAHS to an educational game example. The PAHS framework maps student-content interaction types to determine suitable learning activities in the application domain and maps the activities to design-for-engagement and Merrill's First Principles engagement requirements. Consequently, all these are mapped to Bloom's revised taxonomy to validate the factoring in of incremental cognitive complexity. It is hoped that with more development on the PAHS, it will become a reference model, which can be instantiated in order to inform and validate the design of interactions for both traditional learning and less structured learning in terms of the degree of engagement and that the synergized PAHS will result in an easy means for scaling successful interaction design to reduce error, cost and encourage new innovations.

References

[1] Anderson, L., Krathwohl, D. (eds.): A taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of Educational Objectives. Longman, New York (2001)
[2] Kirkpatrick, D.L.: Evaluating Training Programs: The Four Levels. Berrett- Koehler, San Francisco (1994)
[3] Shortridge, A.: Interactive web-based instruction: What is it? And how can it be achieved? Journal of Instructional Science and Technology 4(1) (March 2001)
[4] Uttendorfer, M.: Interactivity in an online course: Making it more than page turning. In: Proceedings of World Conference on e-Learning in Corporate, Government, Healthcare and Higher Education, pp. 147–149. AACE, Cheasapeake (2003)
[5] Dunlap, J.C., Sobel, D., Sands, D.I.: Designing for deep and meaningful student-to-content interactions. TechTrends 5(4), 20–29 (2007)
[6] Kearsley, G., Shneiderman, B.: Engagement theory: A framework for technology- based based teaching and learning. Educational Technology 38(5), 20–23 (1998)
[7] Preece, J., Rogers, Y., Sharp, H.: Interaction design: Beyond human-computer interaction. John-Wiley, New York (2002)
[8] Merrill, M.D.: First principles of instruction Educational Technology. Research & Development 50(3), 43–59 (2002)
[9] Gamma, E., Helm, R., Johnson, R., Vlissides, J.M.: Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley, New York (2000)
[10] Borchers, J.: A pattern approach to interaction design. Wiley & Sons, New York (2001)
[11] Nielsen, J.: Usability engineering. Morgan Kaufmann, San Francisco (1993)
[12] Pedagogical patterns, http://www.pedagogicalpatterns.org/

Design and Implement of Game Speech Interaction Based on Speech Synthesis Technique

Xujie Wang and Ruwei Yun

Educational Game Research Center of Nanjing Normal University, China

Abstract. Game speech interaction is an fantastic interactive mode but hasn't received enough attention. This study first summarizes features of game speech interaction. And Then, based on speech synthesis technique, we design a speech interaction module which also supported by speech conversion technique. The ultima purpose is to strengthen the interaction between game and player.

Keywords: Speech Synthesis, Speech Interaction, TTS, Speech Conversion.

1 Introduction

With the rapid development of computer and network technology, electronic game has become a sort of main entertainments. Electronic game is also considered "the ninth art", in the same position with music, movie and other traditional art. Furthermore, as creation for business purpose, it has exceed other art because of the accurate simulating of real life. Participation and interactive are unique features of electronic game. Through text, image and sound, electronic game delivers emotional experiences to players. With the accumulation of game time, emotional experiences accumulate either. Sound plays an important role in game but always ignored.

Sound in game is classified into three categories: music, sound effects and speech. Speech is the main element which used to connects player and game. Generally, it contains short voice and dialogue. Using sound in game has purposes for significance and emotion. Significance purpose means that use sound to deliver game information, such as "under attack". Emotion purpose means that use sound to deliver emotional experiences. For example, speech in NPC dialogue is used to direct or adjust players' feeling.

With help of sound, players will immerse into game much easier. Background music is used to enhance ambience, while speech in CG is introducing the scene and scenario to players. A lot of games prove that sound can not only clew game information, but also strengthen the expressive force of game story.

Speech interaction contains speech syntheis and speech recognition. Speech recognition technique is far from mature but speech syntheis has been widely applied. Therefore, this study based on speech synthesis technology and apply it in game speech interaction. The purpose is to strengthen the interaction between game and player.

2 Related Work

Emotioneering™ is the term David Freeman created and it's his registered trade mark. He is the author of "Creating Emotion in Games". David Freeman identifies 32

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 371–380, 2008.

categories of emotioneering techniques and over 300 individual techniques. Many of these are focused on making game elements such as non-player characters, dialog, and plot more interesting or deeper. He also aimed to improve the "chemistry" and relationships between NPCs and between the player and NPCs. Defining interesting and deep groups of NPCs was also addressed, as were group bonding techniques. Most of these techniques use speech as an exhibition.

Stavroula-Evita Fotinea and George Tambouratzis research in Greek Time Domain Speech Synthesis. Their thesis, "A Methodology for Creating a Segment Inventory for Greek Time Domain Speech Synthesis", focuses on the systematic design of a segment database which has been used to support a time-domain speech synthesis system for the Greek language. Emphasis is placed on the comparison of the process-derived corpus to naturally-occurring corpora with respect to their suitability for use in time-domain speech synthesis. The proposed methodology generates a corpus characterised by a near-minimal size and which provides a complete coverage of the Greek language.

Aniruddha Sen and K Samudravijaya are the author of "Indian accent text-to-speech system for web browsing". Incorporation of speech and Indian scripts can greatly enhance the accessibility of web information among common people. This paper describes a "web reader" which "reads out" the textual contents of a selected web page in Hindi or in English with Indian accent. The text-to-speech conversion is performed in three stages: text analysis, to establish pronunciation, phoneme to acoustic–phonetic parameter conversion and, lastly, parameter-to-speech conversion through a production model. Different types of voices are used to read special messages.

Karina Evgrafova is the author of "The Sound Database Formation for the Allophone-Based Model for English Concatenative Speech Synthesis". This paper describes the development of the sound database for the allophone-based model for English concatenative speech synthesis. The procedure of the sound unit inventory construction is described and its main results are presented. At present moment the optimized sound units inventory of the allophonic database for English concatenative speech synthesis contains 1200 elements (1000 vowel allophones and 200 consonant allophones).

Long Qin, Zhen-Hua Ling, Yi-Jian Wu, Bu-Fan Zhang, and Ren-Hua Wang are from iFLYTEK Speech Lab, University of Science and Technology of China, Hefei. In "HMM-Based Emotional Speech Synthesis Using Average Emotion Model", they present a technique for synthesizing emotional speech based on an emotion-independent model which is called "average emotion" model. The average emotion model is trained using a multi-emotion speech database. Applying a MLLR-based model adaptation method, we can transform the average emotion model to present the target emotion which is not included in the training data.

3 Features of Game Speech Interaction

Compared with traditional record play method, speech synthesis takes less memory. It could be more flexible and effective too. Because the file space is reduced, speech can be used in more places and other languages' version can be added. The game developer

only need to prepare actor's lines in text and set the speech engine. Speech can be read out automatically. There's several requests of game speech:

3.1 Natural and Smooth

Animation is used in game to represent polts, connect scenarios etc. Animation is composed by image and sound. Speech is a main part of sound. Even be much vivider, the image is virtual after all. But speech comes from real life, it can involve players into game much easier. Speech must accord with the special requirements of the game, such as plot's rhythm, people's characters etc. It's a necessary condition for virtual space's authenticity. This requires natural and smooth speech.

Fig. 1. Animation in "Warcraft III: Frozen Throne"

3.2 Timely

Interactive is a main feature of game and the feedback of game is rather single now. Use speech to feed back player is more attractive. Different kinds of game have different request: RPG game always use feedback in NPC dialogue but RTS game use feedback for controlled unit.

NPC dialogue is an important technique in RPG game. NPC deeping technique and NPC interest technique in Emotioneering™ often use dialogue and movement as media. The NPC dialogue appears in text on screen isn't intuitionistic enough and could hardly give player emotional experience. We can set characteristic voices for main NPC characters and use speech to give player necessary feedback, immediately.

Player must control units in RTS games fast and careful. For feeding back player's operation, slogan and sound effert should be emitted while unit is enabling. It's very necessary in RTS game. For example, nearly every unit in "Warcraft III: Frozen Throne" has its own slogan. The content of slogan is commonly for inspiriting fight or awaiting orders. If the unit is zoetic or controlled by life-form, slogan is a better choice than sound effert only. Characteristic unit response can inpress player profoundly and enhance the quality of feedback. Player acts general, god and other characters. Speech feeback follows the operation can violently excite player and give them better emotional experience.

- Ready for action!
- Yes my liege!
- Orders?
- Say the word!
- Aye milord!
- On my way!
- ……

Fig. 2. Orcs grunt's speech in "Warcraft III: Frozen Throne"

3.3 Humanized

Traditionally, game will use some sound efferts with text on screen to cluw players. But it's too mechanical and lack of emotion. Humanized speech clew could arose player's emotion experience.

"Warcraft III: Frozen Throne" provides some speech cluw of assistant imformation, such as lack of resource, building completing, under attack etc. It's a common trait of RTS game. Player's game experience should be well advised, and also should accord with game speed. Speech can embodies authenticity, construct ambience and increase mental stimulation. And speech is also a powerful tool used to materialize interactive. Speech can offer timely and effective imformation. It could help ears and eyes get information respectively.

Players may want to set the system personally while playing, such as sound on/off, save mode, screen size etc. System has a default setting and can not give player immediate responses or suggestions. For example, the configuration of a player's computer can not afford high quality graphics effects. If there's no cluw to warn player when a "high" option is choosed, computer may crash or game may terminate. Speech used here to give player appropriate cluw can make the system more friendly and effective.

Fig. 3. Image option in "Warcraft III: Frozen Throne"

4 The Design of Speech Interaction Module

Microsoft, IBM and other companies have developed speech recognition and synthesis engine. Microsoft Speech SDK 5.1 can roundly support English and Chinese speech application's development and it's free. So we use it in the Speech Interaction Module:

4.1 Module Structure

Speech synthesis is an important part of man-machine conversation which has a long research history. Speech synthesis technique classifies three species: ① Synthesis by parameter. Because of the complex arithmetic and information losing in contract, the speech it synthesized could hardly be natural and cleary. ② Synthesis by wave. It is used in the the aspect that don't need to extract parameter in speech synthesis. Through choosing wave in the voice database which synthesized by natural voice, it connects the wave and output. ③ Synthesis by rule. Through controlling preconcerted sign list, it can synthesize any speech.

Microsoft Speech SDK 5.1 comprises correlative groupwares, API, detailed technical information and help documents of the engine. It adopts COM standard in development. Bottom agreements with the form of COM groupwares are independent of application layer. It can help programmers to evade those difficult speech technique and use COM groupwares to accomplish a series work of speech managing. The engine is responsible for speech synthesis, so programmer can focus on the application and use correlative SAPI to implement.

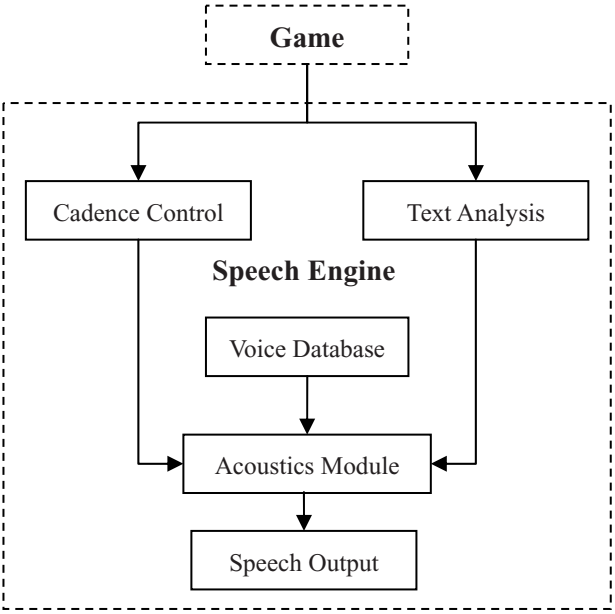


Fig. 4. Speech interaction module

4.2 Speech Conversion

Speech conversion technique is an expanding of speech synthesis technique. Through changing the spectrum, we can make one's voice sounded like another. This technique is very useful in game speech. Recently, the simulatation of source information gets more attention. In games, simulatation of NPC's voice is not only tamber simulatation, but also cadence simulatation. A further research in this field is prosody conversion.

The information of character's speech contains two aspects: source information and track information. Source information comes from shaking of vocal cords and reflects in changes of pitch. It's measured by baseband value. On the other hand, track information comes from the shape of vocal cords. It includes content and voice feature, then reflects in the spectrum distribution. For simulating different features of speakers, speech conversion technique is produced.

Speech conversion is mostly track information conversion, i.e. the conversion of spectrum information. The target is to find rules of mode conversion, which assure the information changeless. The conversed voice should be the same as the source voice and the pronunciation feature changes to the target voice. For getting the conversion rules, we need to record a library of parallel voice materials which speaked by source voice and target voice. Training and conversion are the necessary two approaches.

In the training phase, the voices of source speaker's is compared to the target speaker's and mapping rules which show the relation of the spectrum parameter in two voices will be found. In the conversion phase, the system converse the spectrum feature of source voice with the rules gotten before. After conversion, the target voice will have

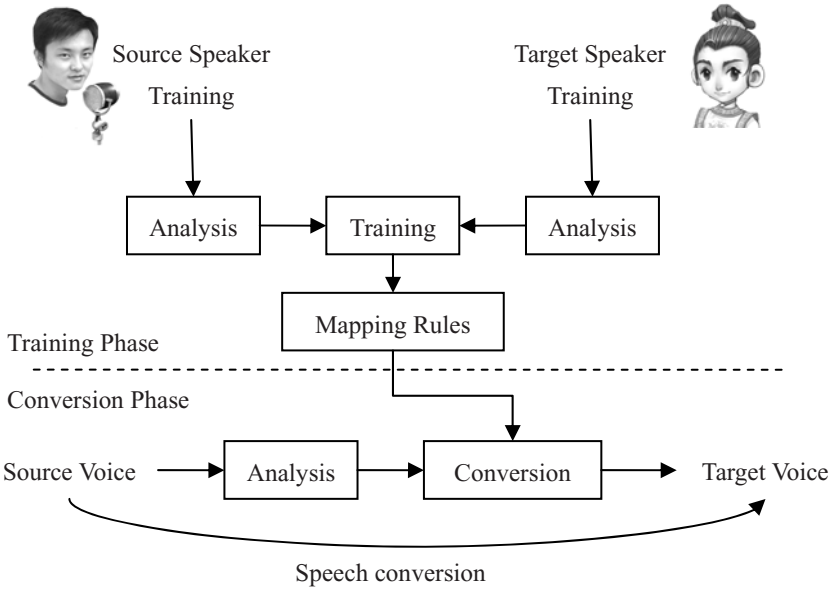


Fig. 5. The process of speech conversion

the feature of source speaker. With speech conversion technique, game developer can set more characteristic voice easier.

5 The Implement of Parameter Adjustable Speech Reading

Microsoft Speech SDK contains API for Text-to-Speech and API for Speech Recognition. Through using API for Text-to-Speech, programmer can easily develop powerful Text-to-Speech application. Speech object library encapsulates details of speech synthesis engine and provides upper interfaces for application to access.

5.1 Instruction of Main Function

The SpVoice object brings the text-to-speech (TTS) engine capabilities to applications using SAPI automation. An application can create numerous SpVoice objects, each independent of and capable of interacting with the others. An SpVoice object, usually referred to simply as a voice, is created with default property settings so that it is ready to speak immediately.

- Attribute

a. Voice: Denote pronunciation style, system will use correlative voice database to read. These default four styles are Microsoft simplified chinese, Microsoft mary, Microsoft mike and Microsoft sam.

b. Rate: The speed of speech . Value range is from -10 to 10.

c. Volume: The volume of speech. Value range is from 0 to 100.

- Method

a. Speak: Accomplish the assignment of transforming text information to speech with the parameters setted. The method has two parameters as test and flags which used to appoint reading text and PRI of mood.

b. Pause: Pause the reading courses in use of the object.

c. Resume: Resume the reading courses in use of the object.

5.2 An Instance

With help of Microsoft Speech SDK, programmer just need to access supported interfaces and leave bottom work to speech synthesis engine.

Speech Application SDK must be installed before programming. The newest version is SAPI 5.1 which have language packages of Chinese, Japanese and English. After being installed, the component "Microsoft voice text" can be used convenient to realize speech read. The main code of an instance programmed by Visual Basic 6.0 is below:

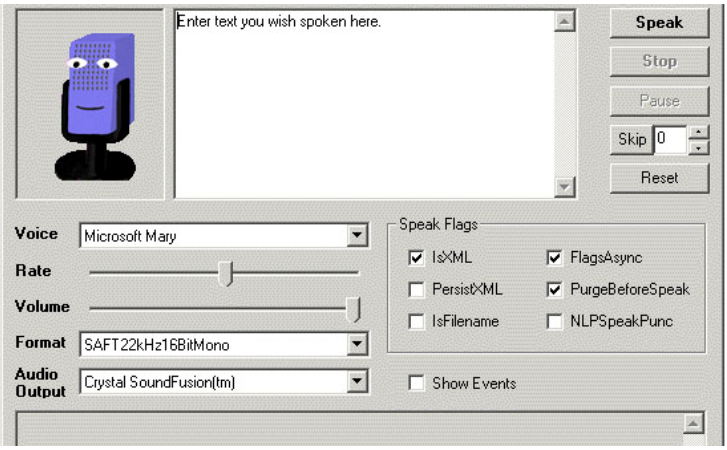


Fig. 6. VB TTS Application

- Common Event

```
'First, declare the main SAPI object we are using in this
sample. It is created inside Form_Load and released inside
Form_Unload.
  Dim WithEvents Voice As SpVoice
  'm_speaking indicates whether a speak task is in progress
  'm_paused indicates whether Voice.Pause is called
  Private m_bSpeaking As Boolean
  Private m_bPaused As Boolean
```

- Form_Load Event

```
'Creates the voice object first
Set Voice = New SpVoice
'Load the voices combo box
Dim Token As ISpeechObjectToken
For Each Token In Voice.GetVoices
VoiceCB.AddItem (Token.GetDescription())
Next
VoiceCB.ListIndex = 0
'set rate and volume to the same as the Voice
RateSldr.Value = Voice.Rate
VolumeSldr.Value = Voice.Volume
```

- Speak Button Event

```
'If it's paused and some text still remains to be spoken, Speak button acts the same as Resume button. However a programmer can choose to speak from the beginning again or any other behavior.
'In other cases, we speak the text with given flags.
If Not (m_bPaused And m_bSpeaking) Then
    'just speak the text with the given flags
    Voice.Speak MainTxtBox.Text, m_speakFlags
End If
```

- Rate Slider Event

```
Private Sub RateSldr_Scroll()
    Voice.Rate = RateSldr.Value
End Sub
```

- Voice ComboBox Event

```
Private Sub VoiceCB_Click()
    ' change the voice to the selected one
    Set Voice.Voice =
     Voice.GetVoices().Item(VoiceCB.ListIndex)
End Sub
```

6 Conclusion and Ongoing Work

Game interaction design is a developping field and many techniques in design need speech to reflect. Whereas speech didn't get enough regard as an important interaction. The thesis discussed the application of speech synthesis technique in game interaction design and gave an example of speech synthesis application.

Isolated technique or isolated design is powerless unless they are well integrated. Therefore the thesis lucubrated requests of speech in game, summarized several speech

modes in game. The probability of changing traditional sound files mode to speech synthesis mode were also discussed.

Game asks for high quality speeches, so the key of application is natural and effective synthesized speech. How to improve the quality of synthesized speech will become an important part in furture research. At the same time, other in-depth speech synthesis technique, such as emotional speech synthesis, will be researched. The application of speech in game will get more attention and play a more important role in game.

References

1. Freeman, D.: Creating Emotion in Games, New Riders Games (September 2003) ISBN: 1592730078
2. Fotinea, S.-L., Tambouratzis, G.: A Methodology for Creating a Segment Inventory for Greek Time Domain Speech Synthesis. International Journal of Speech Technology (June 2005), doi:10.1007/s10772-005-2167-5
3. Sen, A., Samudravijaya, K.: Indian accent text-to-speech system for web browsing, Sadhana (February 2002), doi: 10.1007/BF02703316
4. Evgrafova, K.: The Sound Database Formation for the Allophone-Based Model for Eng-lish Concatenative Speech Synthesis. Text, Speech and Dialogue, doi:10.1007/11551874_28
5. Qin, L., Ling, Z.-H., Wu, Y.-J., Zhang, B.-F., Wang, R.-H.: HMM-Based Emo-tional Speech Synthesis Using Average Emotion Model, Chinese Spoken Language Processing, doi: 10.1007/11939993_27
6. Black, A., Campbell, N.: Optimising selection of units from speech databases for concatenative synthesis. In: Proceedings of Eurospeech (1995)
7. Michael Macon, W., Mark Clements, A.: An Enhanced ABS/OLA Sinusoidal Model For Waveform Synthesis In TTS. In: Proceedings of Eurospeech (September 1999)
8. Bailly, G., Campbell, N.: ISCA Special Session: Hot top ics in speech synthesis. In: Proceedings of the European Conference on Speech Communication and Technology (2003)

Two-Arm Haptic Force-Feedbacked Aid for the Shoulder and Elbow Telerehabilitation

Patrick Salamin[1], Daniel Thalmann[1], Frédéric Vexo[1], and Stéphanie Giroud[2]

[1] VRLab -EPFL, Switzerland
http://vrlab.epfl.ch
[2] Hôpital du Chablais
http://www.hopitalduchablais.ch

Abstract. In this paper we present a telerehabilitation system aiming to help physiotherapists on the shoulder and elbow treatment. Our system is based on a two-arm haptic force feedback to avoid excessive efforts and discomfort with the spinal column and is remotely controlled by smart phone. The validation of our system, with the help of muscular effort measurements (EMG) and supervised by a physiotherapist, provides very promising results.

1 Introduction

Patients suffering from muscle or ligament diseases need training for their rehabilitation after the medical diagnostic. This kind of treatment can be applied only in a hospital or in a physiotherapist offce. As mentioned in [2], having a machine able to handle the patients in (even rural) hospitals would then be very useful, but there is still a need of a physiotherapist during the treatment. If the patients want to get the best rehabilitation, they need an expert physiotherapist for their injured limb. Unfortunately, these physiotherapists are most of the times working in the metropolis because of their popularity. It would thus be a strong asset if they could remotely control these machine from their offce. With such a system, the physiotherapists would easily train to better rehabilitate the patients almost at home.

In this paper, we propose a remotely controlled system for patient shoulders and elbows rehabilitation and training of physiotherapists. We decide to work with the Immersion[1] Haptic Workstation TM instead of other VR haptic devices. In fact, it is a two-armed-based system [11] that allows a well-balanced effort of the patients between their right and left limbs. This avoids excessive torsions and efforts with the spinal column. Moreover, this machine can be easily tele-operated by a smartphone via Internet in contrary of other "home medical systems" like the Biodex[2] or the Cybex[3] that are not less cumbersome.

We first present a short overview of the related fields that lead us to this improved system. Secondly, we describe the overall system and its implementation. The third

[1] http://www.sensable.com
[2] http://www.biodex.com/rehab/rehab.htm
[3] http://ecybex.com

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 381–390, 2008.

part is dedicated to the system validation by a physiotherapist with the help of an ElectroMyoGraph (EMG) to measure the muscular effort of the patient. Finally, we conclude with the advantages of our system and its future possible improvements.

2 Related Works

Using machines as an aid for the rehabilitation is not a new concept. Researchers already thought about it in 1965 [13] but only the veterans were targeted till 1967[7]. In the following years, researches tends to aid handicapped people to substitute their "damaged" limbs [4].

Since the beginning of this millennium, rehabilitation became a very fashion topic. An overview of the current machines aiding to rehabilitation can be found in [16]. Researchers now also work on solutions for incapacitated people in order to help them to recover the use of their limbs that we split into two parts: the lower upper limbs. These last ones can be divided in four main parts that have to be considered: shoulder, elbow, wrist and fingers. Tsagarakis et al. [14] developed one of the first machines aiding to people rehabilitation. Unfortunately, the applied forces were bounded to 2kgs, which limited a lot the efficiency of the prototype for a complete treatment.

David Jack et al. focused on the fingers rehabilitation with the help of force-feedback gloves [6]. Others researchers made complementary works and developed a 6DOF machine for the shoulder, elbow and wrist rehabilitation for one hand at a time [3]. Nevertheless, all these devices require the presence of a doctor during the patient healing. In 2005, Demiris et al. highlighted in [2] the interest to develop a machine that would allow healing the patient at distance. They involved by this way the telerehabilitation in the scientific research for the patients' well-being.

Based on these researches and with the help of a physiotherapist, we assumed that the HW would be very efficient and promising to treat upper limbs injuries like the shoulders and the elbows. We present our system in the following chapter.

3 System Description and Implementation

We present hereinafter our application and the artifacts that should improve its believability and efficiency.

3.1 System Architecture

As the goal of this application is the telerehabilitation, we want to allow the physiotherapists to be completely mobile and independent from the patient location. A PDA with an integrated webcam (e.g. a PDA-phone) seems thus to perfectly fit to our requirements. By this way, the patient can see a video-streaming of the physiotherapist sent by the PDA during its treatment. The physiotherapist is informed of the patient arms location with a 3D interface created through the MVisio 3D graphics engine [12] and the data sent by the HW. As you can see on the first picture from the left on the Figure 1, the physiotherapist does not need to be with the patient or in his/her office to begin the treatment.

The patient is wearing a HMD while being seated in the HW as shown on the third picture from the left on the Figure 1. The HW allows the therapist to apply different forces in a range of 0 gram to 10k grams in every direction on each arm independently. As we can see in [15], the shoulder can be injured in several ways. It is thus useful to be able to apply accurate forces in order to help him/her to move his/her arm at the beginning and then to apply a force against his/her movement. As the HW applies the forces on the wrist and as we want the user to move his/her shoulder, we hold the patient arm straight with the help of a harness. The main advantage of using the HW – even if it is very expensive – resides in the possibility to apply forces on both arms. By this way, if we apply symmetrical forces on both arms, the patient will not try to twist his/her trunk to execute the movement. The rehabilitation would then not lead to new problems with the spinal column.

Moreover, as we are working with the HW, with the help of the paradigm developed by Renaud Ott in [10], it is possible to compensate the gravity effect as if the patient arm was in weightlessness. The patients can be treated from the beginning with micro-gravity to their complete recovery with forces around 10kgs. It is also very interesting for the physiotherapist to be able to apply a constant and exact force (less than one gram). This would avoid most of errors due to human factor (deviance of the applied force that is neither exact nor constant).

Finally, electrodes are applied on the patient skin to detect the muscles activity with the help of an EMG. The provided information is sent to the physiotherapist who can better evaluate and appreciate the force to apply on the patient and the evolution of his/her rehabilitation.

3.2 Improvements for the User Immersion

The physiotherapist only needs to have a webcam and a PDA with a very simple interface as you can see on the second picture from the left on the Figure 1. He/she can see the current patient position represented by an avatar. The doctor can also easily change the forces applied on the patient with the help of the stylus by indicating the concerned wrist and the force direction and amplitude which depends on the line he/she draws (values in grams written on the screen).

On the patient side of our application, we know by experience that VR systems are quite invasive and can stress the user. The simulation would then be less efficient, traumatizing and even harmful for the patient. This is why the user is immerged in a virtual environment with a relaxing landscape like a beach or mountains, depending on his/her preference as you can see on the right picture of the Figure 1. A sweet relaxing music as background noise also contribute to strongly reduce the patient anxiety due to the VR machines [8].

Moreover, the patient can see a realistic representation of his/her hands which follows in real-time the real position and orientation of the real ones. Notice that the force is applied on the user wrist(s) is represented as for puppets: the concerned wrist is caught by a wire. With this artifact, the patient can see (right picture of the Figure 1 in which direction the force is applied and also its intensity (a second wider red wire indicates it). All these artifacts improve the immersion of the patient who could be seriously perturbed by the used hardware that we present in the following section.

Once done, some scenarios could be easily added to the simulation. For example, the user could have to touch a virtual ball the physiotherapist would move around specific places. It has already been proved that such a playful simulation improves rehabilitation results [1][9].

Finally, during the simulation, the patient has a "Window to the World" which allows him/her to see the physiotherapist and to listen to him/her during the session. As shown in [5], it is very relaxing for the patient to have a multimodal link with the doctor while being in the virtual environment.

In order to prove the efficiency of our system we made some tests that are presented in the following section.

4 Experiments

We display a relaxing landscape during the complete session. Moreover, we first let the patient in the environment with a sweet music in background during fifteen minutes in order to compensate the stress possibly brought by the VR engines. Once the user seems to be relaxed, we begin the treatment. We decide to call "neutral position" the position in which a tester tighten ones' arms close to one's trunk on the vertical plane without moving.

Obviously, depending on his/her recovery status, a help (or a constraint) is applied by the HW when the patient tries to perform the asked movements shown on the Figure 2. We have to remind that for every exercise, the patient moves both arms in a synchronous and parallel way (or symmetrical when they are on the same plane, e.g. for the abduction we present hereinafter) to avoid useless and dangerous efforts and torsions on the spinal column.

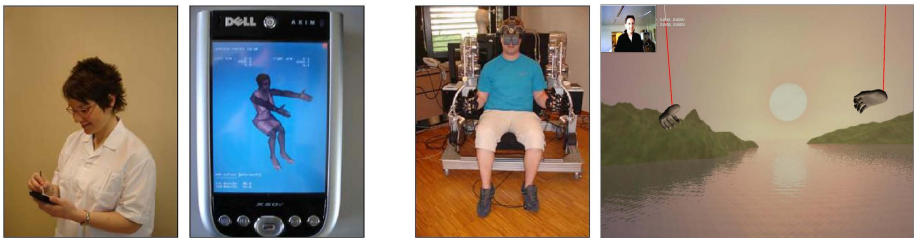


Fig. 1. On the left, the physiotherapist (hardware and interface); on the right, the patient (hardware and interface) during the rehabilitation exercises

We present hereinafter four exercises for the shoulder rehabilitation: inflection, abduction, lateral and medial rotation. Eight testers (two females) participate to the simulation to check the efficiency and the limitations of our system, but unfortunately (or hopefully) no one had any shoulder or elbow troubles.

The inflection exercise consists in moving ones arm in a vertical plane ahead of oneself as shown on the left pictures of the Figure 2. The concerned muscles are mainly the anterior deltoid and the pectoralis major, but the coraco brachial and biceps are also

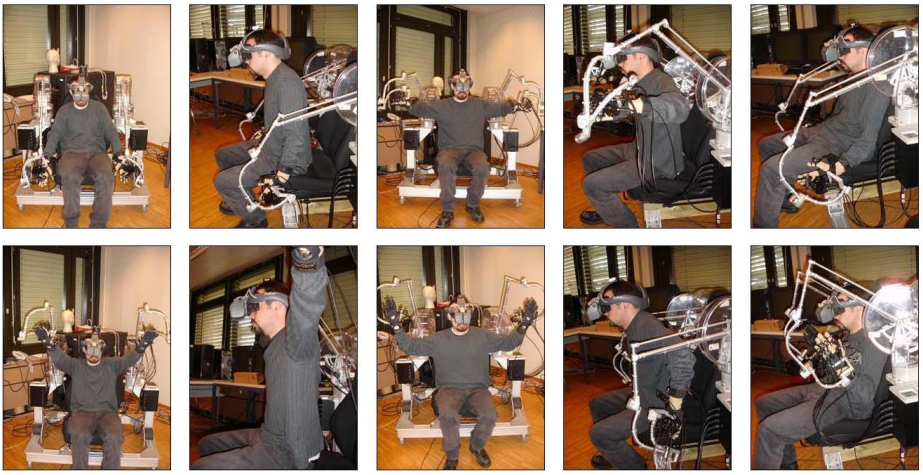


Fig. 2. Exercises for the rehabilitation (from the left to the right): inflection, abduction, lateral rotation, medial rotation for the shoulder and elbow inflection

working during this exercise. In the following section, you will see that we put electrodes on the main-working muscles to check their activity during the tests.

In the abduction exercise, as you can see on the second pictures from the left of the Figure 2, the patient must move his/her arms in a vertical plane on his/her sides (instead of ahead). In this experiment, the most acting muscles are the middle deltoid and the supraspinatus, whose behavior will be observed and commented in the following section.

Concerning the rotation exercises, the patient arms start in position shown on the top third and fourth pictures from the left in the Figure 2. For the lateral rotation, the forearms raise up to the vertical. In this case, the main working muscles are the posterior deltoid, the teres mino and infraspinatus which are working together (bottom centered picture). For the medial rotation, the forearms fall down (almost) to the vertical. The main detectable activated muscles are the anterior deltoid and the pectoralis major but some others like the subscapularis (under the pectoralis major), the latissimus dorsi and the teres major are also working but less perceptible with the EMG.

Concerning the elbow rehabilitation (right pictures in the Figure 2, we first assume the patient keeps the arm in the "neutral position" (close to the body trunk). After this, he/she must move vertically in a synchronous way the forearms from the bottom to the top and vice versa. The most important muscles used to perform this exercise are the biceps and the triceps on which we also put electrodes in order to obtain their activity. We present in the next section typical results we obtained during our sessions with the help of our system and an EMG.

5 Discussion of Results

As written before, in order to avoid bias due to the VR devices used for these experiments, the user is first seating in the HW with sweet music and a relaxing landscape

during fifteen minutes. He/she performs then twice the exercise, e.g. “down – up – down – up – down” for the shoulder inflection. We present hereinafter the results we obtained but, in order to better understand the provided graphs, we first define two terms concerning the muscles:

– effectors: These muscles are used to execute a movement. They give the possibility to the patient to accomplish a movement. If, in our case, the shoulder does not move, they do not act. Their EMG representation has a big difference when the shoulder moves or not.

– stabilizers: These muscles are always in activity, even when the shoulder does not move because they have to stabilize the joint to avoid (e.g. a dislocation due to the gravity effect). They almost do not act to perform a movement. Their EMG representation seems to be almost constant.

For the first exercise we present, the inflection, the electrodes are located on the anterior deltoid and the pectoralis major. They are respectively represented on the EMG graphs by the red and the blue lines on the Figure 3. Notice that we can perfectly see the role of those muscles: both are effector muscles and work during the whole movement (up -stay -down) while the patient arm is not in the “neutral position”. We can also notice that the main used muscle for the inflection is the anterior deltoid while the pectoralis major only helps it for this action. It is also proved on the right graph of the right picture of the Figure 3 because when a force is applied against the patient movement, the pectoralis major acts during the complete movement.

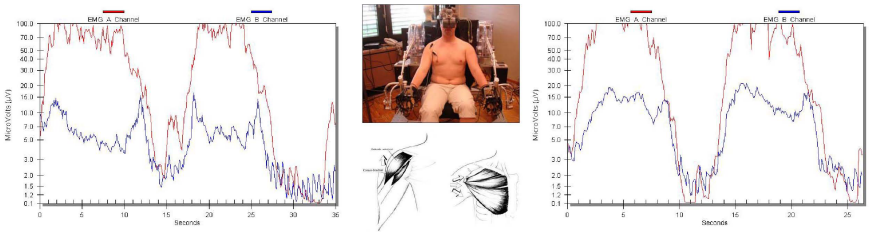


Fig. 3. Inflection exercise for the shoulder: (left)with help, (right)countered to and (center) position of the electrodes (anterior deltoid on the bottom left and pectoralis major on the right)

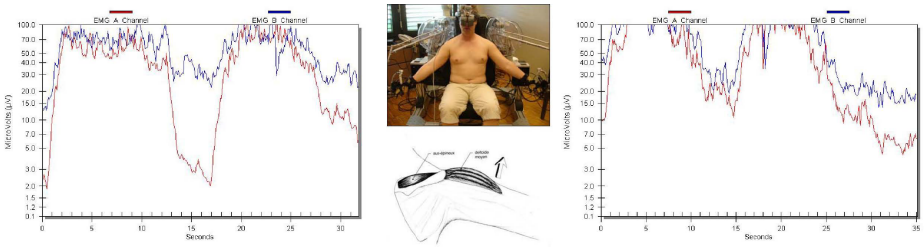


Fig. 4. Abduction exercise for the shoulder: (left) with help, (right) countered to and (center) position of the electrodes (supraspinatus on the bottom left and middle deltoid on the right)

The abduction exercise mainly involves the activity of the middle deltoid (in red) and the supraspinatus (in blue). Their activity during the experiment can be seen on the Figure 4. Once again, we decided to check the activity of an effector (the deltoid) and a stabilizer (the supraspinatus) to verify if the HW is really efficient when it helps or counter to the patient movements. We can see on the Figure 4 that the muscles, during the movement, are used a lot – even with the help provided by the HW. But we can also remark that they are almost not used when the user is in the "neutral position".

For the lateral rotation (Figure 5), the electrodes are located on the posterior deltoid (blue line) and both teres minor and infraspinatus because they are linked(red line). As those last ones are stabilizer muscles, they seem to be always active during the experiment while the posterior deltoid can "have a rest" when the patient is in the "neutral position". In this case, the main difference of muscle activity appears for the stabilizer ones because the applied countering forces are strong (right graph of the Figure 5).

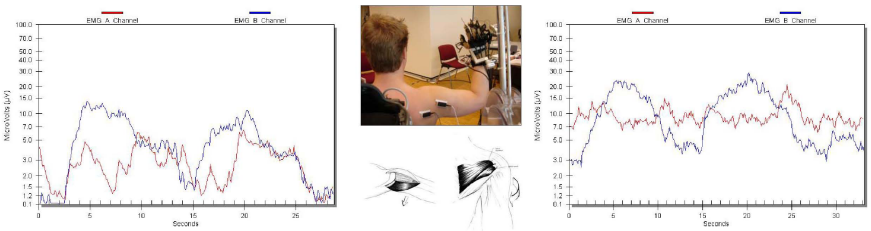


Fig. 5. Lateral rotation exercise for the shoulder: (left) with help, (right) countered to and (center) position of the electrodes (posterior deltoid on the bottom left, teres minor and infraspinatus on the right)

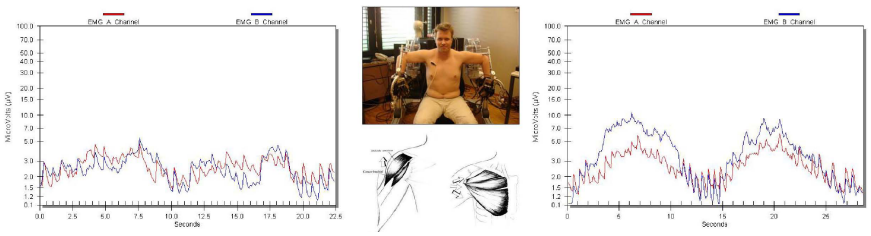


Fig. 6. Medial rotation exercise for the shoulder: (left) with help, (right) countered to and (center) position of the electrodes (anterior deltoid on the bottom left, pectoralis major and subscapularis (hidden) on the right)

The medial rotation (Figure 6) involves lots of muscles. The anterior deltoid the pectoralis major are the most active and interesting to analyze with an EMG. Among the others, we can cite the subscapularis, the altissimus dorsi and the teres major. As we can see it on the left and right pictures, both analyzed muscles are effector ones (only the subscapuralis is a stabilizer). They thus almost follow the same curve and we can see that the forces applied by the HW are also quite efficient for this exercise. Notice that for this exercise, the applied counter forces are pointing ahead and up.

The last exercise we present in this paper concerns the elbow rehabilitation (Figure 7). We ask the patient to perform elbow inflections and we measure the biceps (red line) and triceps (blue line) activity. It is interesting to see that the triceps, which is normally an effector muscle for the extension of the elbow, also acts in this exercise and almost follow the biceps curve. Otherwise, the efficiency of the HW and the applied force is obvious. A big difference between the muscle activity between the left (helping forces) and the right graphs where a force is applied to the front and to the bottom can be noticed.

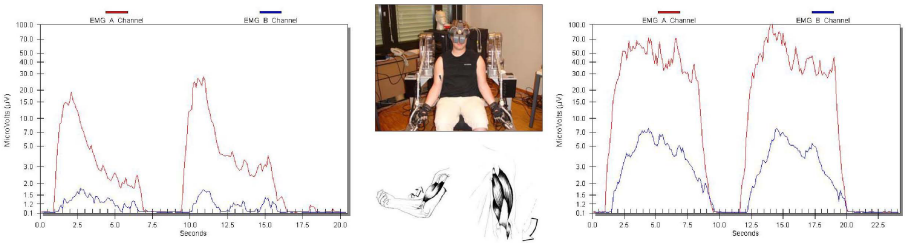


Fig. 7. Inflection exercise for the elbow: (left) with help, (right) countered to and (center) position of the electrodes (biceps on the bottom left and triceps on the right)

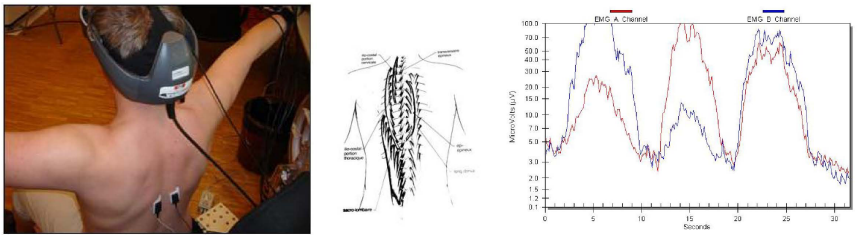


Fig. 8. Back muscles difference

Finally, we prove the benefits of working with both arms at the same time. We have made this assumption because when a physiotherapist asks a patient to do rehabilitation exercises at home, he/she often propose to perform them with both hands. We then check the activity of the errector spinea (shown in the center of the Figure 8) for the abduction exercise we have already presented. We locate an electrode on the left side (red line of the graph below) of the spinal column and another one on the right side (blue line). In order to obtain the graph on the Figure 8, we first asked the patient to stick up his/her arm and then to lower it. After this, he/she performs the same movement with the other arm and finally with both together. As you can see below, there is a very big difference between the right and left spinal muscles activity when the patient only sticks up one arm. This leads to spinal column torsion and often to backache. This graph prove then the importance to work with both arms during the rehabilitation.

6 Conclusion

The goal in this paper was to present an efficient aid for the shoulder and elbow telerehabilitation. Our application full-fills the tele-operation side that provides obvious advantages for the patients and the physiotherapists training. Sometimes an adaptation time of five to ten minutes was needed to discover the VR material, but none of the patients were really perturbed by them during the simulation. Furthermore, the obtained results seem to prove the efficiency of our system for the patients during all the rehabilitation phase. We can e.g. see in the graphs that a very light force is needed to perform the action when the patient starts his/her rehabilitation. And when he/she has almost recovered all his/her faculties, the HW can apply strong enough forces on his/her arms to finish correctly the rehabilitation. The EMG last graphs also support the idea of the minimal spinal column torsions mandatory for the patient comfort. Moreover, the possibility to cure the patients at distance also really interested our physiotherapist. This technology extends the coverage of this kind of therapy because the patients can be treated in any hospital even if the therapist is not physically present.

Concerning the future works, it could be really interesting, e.g. for the lateral and medial rotation, to provide a support for the elbow during the therapy. The physiotherapist should also take into account an additional five to ten minutes to install the patient in the HW. However this task can be performed by any member of the medical staff.

Acknowledgments

This work has been partially founded by the EU IST-INTUITION Network Of Excellence.

References

1. Betker, A.L., Szturm, T., Moussavi, Z.K., Nett, C.: Game based interactive motivating tool for rehabilitation movements: case study. In: ACST 2006: Proceedings of the 2nd IASTED international conference on Advances in computer science and technology, pp. 157–162. ACTA Press, Anaheim (2006)
2. Demiris, G., Shigaki, C.L., Schopp, L.H.: An evaluation framework for a rural home-based telerehabilitation network. J. Med. Syst. 29(6), 595–603 (2005)
3. Furusho, J., Li, C., Hu, X., Shichi, N., Kikuchi, T., Inoue, A., Nakayama, K., Yamaguchi, Y., Ryu, U.: Development of a 6-dof force display system using er actuators with high-safety. In: VRCIA 2006: Proceedings of the 2006 ACM international conference on Virtual reality con-tinuum and its applications, pp. 405–408. ACM Press, New York (2006)
4. Greenleaf, W.J.: Applying vr to physical medicine and rehabilitation. Commun. ACM 40(8), 42–46 (1997)
5. Gutiérrez, M., Lemoine, P., Thalmann, D., Vexo, F.: Telerehabilitation: controlling haptic virtual environments through handheld interfaces. In: VRST 2004: Proceedings of the ACM symposium on Virtual reality software and technology, pp. 195–200. ACM Press, New York (2004)

6. Jack, D., Boian, R., Merians, A., Adamovich, S.V., Tremaine, M., Recce, M., Burdea, G.C., Poizner, H.: A virtual reality-based exercise program for stroke rehabilitation. In: Assets 2000: Proceedings of the fourth international ACM conference on Assistive technologies, pp. 56–63. ACM Press, New York (2000)
7. Jaffe, D.L.: An overview of programs and projects at the rehabilitation research and development center. In: Assets 1994: Proceedings of the first annual ACM conference on Assistive technologies, pp. 69–76. ACM Press, New York (1994)
8. Kallinen, K.: The effects of background music on using a pocket computer in a cafeteria: im-mersion, emotional responses, and social richness of medium. In: CHI 2004: CHI 2004 extended abstracts on Human factors in computing systems, pp. 1227–1230. ACM Press, New York (2004)
9. Lathan, C., Vice, J.M., Tracey, M., Plaisant, C., Druin, A., Edward, K., Montemayor, J.: Therapeutic play with a storytelling robot. In: CHI 2001: CHI 2001 extended abstracts on Human factors in computing systems, pp. 27–28. ACM, New York (2001)
10. Ott, R., Gutierrez, M., Thalmann, D., Vexo, F.: Improving user comfort in haptic virtual environments through gravity compensation. In: WHC 2005: Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, pp. 401–409. IEEE Computer Society Press, Washington, DC (2005)
11. Ott, R., Perrot, V.D., Thalmann, D., Vexo, F.: Mhaptic: a haptic manipulation library for generic virtual environments. cw 0, 338–345 (2007)
12. Peternier, A., Thalmann, D., Vexo, F.: Mental Vision: A Computer Graphics Teaching Platform. In: Pan, Z., Aylett, R.S., Diener, H., Jin, X., Göbel, S., Li, L. (eds.) Edutainment 2006. LNCS, vol. 3942, pp. 223–232. Springer, Heidelberg (2006)
13. Rose, G., Woodbury, M.A., Fergusson, E.S.: Recent study of teaching: machine aids in rehabilitation therapy. In: Proceedings of the 1965 20th national conference, p. 7. ACM Press, New York (1965), Turner, L.R. (Chairman)
14. Tsagarakis, N.G., Caldwell, D.G.: Development and control of a “soft-actuated” exoskeleton for use in physiotherapy and training. Auton. Robots 15(1), 21–33 (2003)
15. Uhthoff, D.H.K.: Shoulder injury and disability. Technical report, Orthopaedic Surgeon (2002)
16. Yeh, S.-C., Rizzo, A., Zhu, W., Stewart, J., McLaughlin, M., Cohen, I., Jung, Y., Peng, W.: An integrated system: virtual reality, haptics and modern sensing technique (vhs) for post-stroke rehabilitation. In: VRST 2005: Proceedings of the ACM symposium on Virtual reality software and technology, pp. 59–62. ACM, New York (2005)

Vision Based Pose Recognition in Video Game

Dong Heon Jang, Xiang Hua Jin, and TaeYong Kim

Department of Image Engineering, Graduate School of Advanced Imaging Science Multimedia, and Film, Chung-Ang University,
221 Heuseok-Dong Dongjak-gu, 156-756 Seoul, South Korea
tellamon@gmail.com, hyanghwa_kim@naver.com, kimty@cau.ac.kr

Abstract. We present a vision based HCI system which exploits background subtraction comparing local orientation histograms. As a new virtual input device for game control, we focus on extracting coarse pose of the foreground object and its application to video game. The captured image is divided into the cells where the local orientation histogram with Gaussian kernel is computed and compared with the corresponding one using Bhattacharyya distance measure. The orientation histogram-based method is partially robust against illumination change and small moving objects in background. We also propose a vision-based interfacing system to existing game engines and appropriate modules that includes recognition process using neural network. The real-time 3D video games are implemented as a test-bed with the proposed system to prove the presented vision based system is highly applicable to let users control virtual environment without any hard-wired input devices.

Keywords: Vision–based game interface, HCI system, Background subtraction, Orientation histogram, Integral histogram.

1 Introduction

The recent video gaming consoles operate with the motion sensor technology for immersive game playing experience appealing a wide range of game users. The depth-sensing camera is also highly applicable but need extra cost for the end user [10]. Although the vision-based interface is considered as an effective way to capture user inputs without any additional burden such as a data glove, it lacks its robustness in illumination changes and small perturbations such as slightly moving objects in background. In addition, the minimum computation cost is also required for a game application in real-time [6].

The background subtraction methods for segmenting foreground regions are attempted using color or intensity thresholding. To model the variances of background, C. Stauer used mixture of Gaussian on changes of pixel intensity taking several frames to converge each Gaussian [8]. But it needs few frames at initialization step. Recently, a few methods of temporal such as Optical flow and color co-occurrence were introduced for object segmentation in motion [9], [13], but it is infeasible for real-time processing.

In contrast to tracking based algorithm that localizes regions or points [6], [8], [11], in this paper, we analyze a whole image using local histograms despite their previous

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 391–400, 2008.

location in a frame-by frame basis so that the unexpected convergence to local minima can be avoided. Histogram-based object tracking is practiced in many previous researches [2], [11], [14]. A color histogram is easy to compute and is partially robust against small perturbations since it represents color distribution of the area while lacking its spatial information [4], [5]. But, color histogram-based approach fails under illumination variance caused by self-shadow casting foreground object or automatic bright adjustment of camera device. To overcome this problem, using only the color information is a possible solution like Hue channel in HSV color space [1], but it also fails under low brightness because the Hue value isn't defined.

Orientation histogram has been identified as a potential of using histogram-based descriptors for visual tracking. It shows strong performance against illumination changes which cannot be accomplished by color or intensity-based histogram. Moreover, we adopt the local representation of image that describe the background and the foreground by splitting each image into small cells which gives coarse spatial information. For fast computation, the use of the integral histogram [3] and efficient comparing method [1] are exploited and improved to work with orientation-based local histogram.

The extracted foreground cells are used as an input to the neural network where the predefined poses are already trained. Including recognition module, we propose a vision-based interface system whose main purpose is to be seamlessly attached to the existing game engine under given game context and corresponding input map. We have implemented real-time game with vision interface system to show functional efficiency of proposed system for controlling virtual environment. The overall control flow is shown in Figure 1.

The following section explains how to calculate the local orientation histograms with various. Section 3 introduces the vision interface system including the neural network for pose recognition. Section 4 is experiment result with real-time 3D game implementation and the final section concludes the paper.

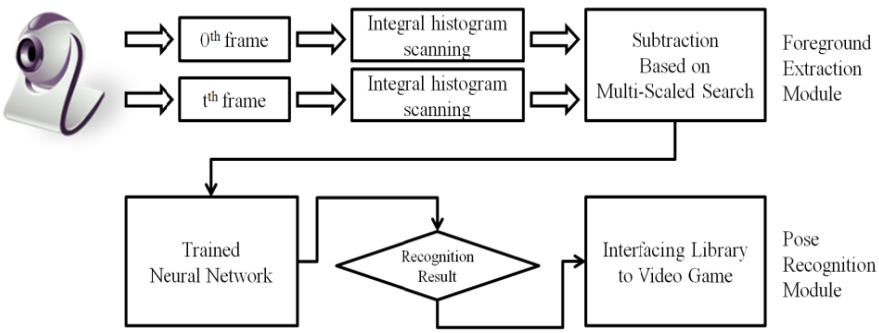


Fig. 1. The overall structure of system

2 Local Orientation Histogram

For the real-time application, it is necessary to reduce the amount of data by grouping neighboring pixels to local region (hereby cell) and by quantizing the feature space

before histogram computation. We divide $V \times H$ sized high-resolution image into $v \times h$ number of cells. And at each cell in image, an N -bin histogram of its neighborhood is computed. In discrete histogram, each bin covers $180/N$ degree of gradient, so we need to choose proper N to trade off quantization error against memory usage. Determination of the number of histogram bins is an important yet unresolved problem in color-based object tracking, so the bin number is empirically set(N =8 in our case) and selection of bin number accounting for environment changes is left for future work.

2.1 Gradient Orientation Histogram

The necessary steps to compute gradient orientation histogram in cell are as follows. Firstly, the gradient of I is computed at each point (x, y) , given $dy = I(x, y+1) - I(x, y-1)$ and $dx = I(x+1, y) - I(x-1, y)$, then the magnitude and the orientation is calculated as follows.

$$m(x, y) = sqrt(dx^2 + dy^2) \text{ , and } \theta(x, y) = \arctan(dy / dx) \tag{1}$$

Secondly, θ is quantized into N bins. The running sum of each bin is computed separately. In order to reduce the effect of noise, the contribution of each point in $\theta(x, y)$ to the corresponding bin is weighted by its magnitude $\mathrm{m}(x, y)$. We also apply Gaussian kernel on 1D histogram for the $\theta(x, y)$ to contribute several bins according to Gaussian weights since limiting the corresponding histogram bin of pixel (x, y) to a single bin is a major reason of quantization errors. The example of gradient orientation histogram is shown in Figure 2. In the next section, we describe how the calculation of histogram can be accelerated by integral histogram and multi-scaled search algorithm.

2.2 Integral Histograms and Subtraction Using Multi-scaled Search

As proposed in [1], the histogram based multi-scaled search algorithm requires multiple extractions of histograms from multiple rectangular cells. The tool enabling this to be done in real time is the integral histogram described in [3].

The integral histogram method is an extension of the integral image data structure described in [7]. The integral image holds the sum of all the pixels contained in the rectangular region at the point (x, y) in the image. This image allows computing the sum for the pixels on arbitrary rectangular regions by considering the 4 integral image values at the corners of the region i.e it is computed in constant time independent of the size of the region. The integral histogram at each position (x, y) is calculated with wavefront scan propagation that 3 neighbor integral histograms and the current orientation value are used. It is done at each frame and consumes the most of computation time and the Figure 3 gives detailed explanation of scanning process.

Then by accessing these integral histograms we can immediately compute the local histogram in a given region with integral data at four corners as follows.

$$H_l = I(right, bottom) - I(right, top) - I(left, bottom) + I(left, top) \tag{2}$$

Given the local histograms of the background and the foreground, we can determine which areas of the current frame contain foreground objects by computing distance of

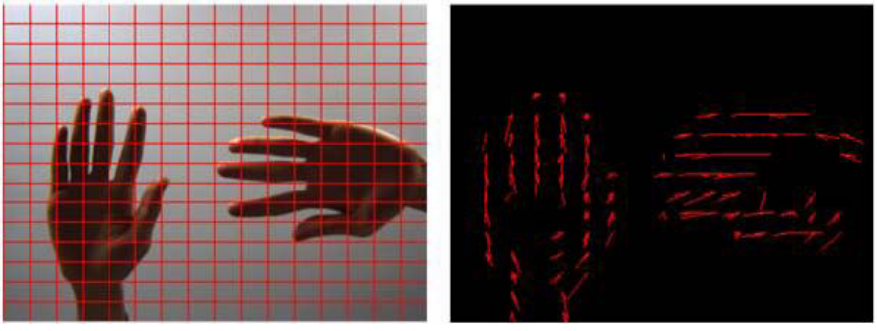


Fig. 2. The test image (left) and its local orientation histogram (right), In histogram plotting, the line direction represents its bin position (angle) and its length is magnitude.

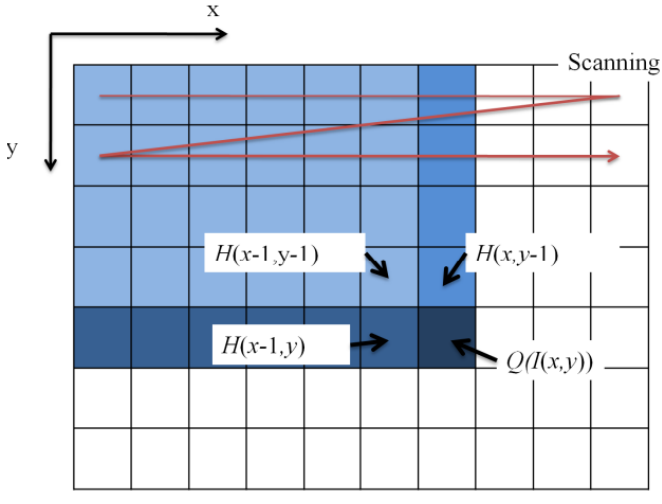


Fig. 3. Integral data at $I(x, y)$ contains the cumulative orientation values from the starting point $(0,0)$ to $I(x, y)$ in rectangular region. The scanning must be done before histogram extraction step. In the wavefront scanning, $H(x, y)$ can be calculated from 3 memory access from $H(x-1, y-1)$, $H(x, y-1)$ and $H(x-1, y)$ and add $Q(I(x-1, y))$.

two histograms. To compare the current local histogram Hlcur with the reference local histogram Hlref at the same cell l, we first normalize the histograms and apply Bhattacharyya distance measurement.

Rather than subtracting cell by cell, we adopt multi-scale search algorithm to skip large region where no foreground objects come in [1]. The key algorithm is that it recursively check histogram distance between foreground and background until the search level reaches to maximum level. Note that there is no additional histogram quantization process for searching at each level. With the help of the pre-scanned integral histogram, multiple extractions of histogram over different scaled regions are done in constant time. This approach mainly has two advantages. The first is that it

suppresses the false positives due to various noises in the background. Another benefit of using the multi-scaled approach is the superior computation speed. It automatically skips the large background area and go deeper level to find foreground cells while cell by cell comparison always consume v×h computation time.

Figure 4 shows the efficiency of multi-scaled search approach. Note that the searched level is highlighted in red rectangles. The large local area implies that it skips the sub region searches which results in suppression of the false positives and computation advantage compared to the cell by cell comparison method. Also the figure shows that the orientation histogram has a feature of illumination invariance.

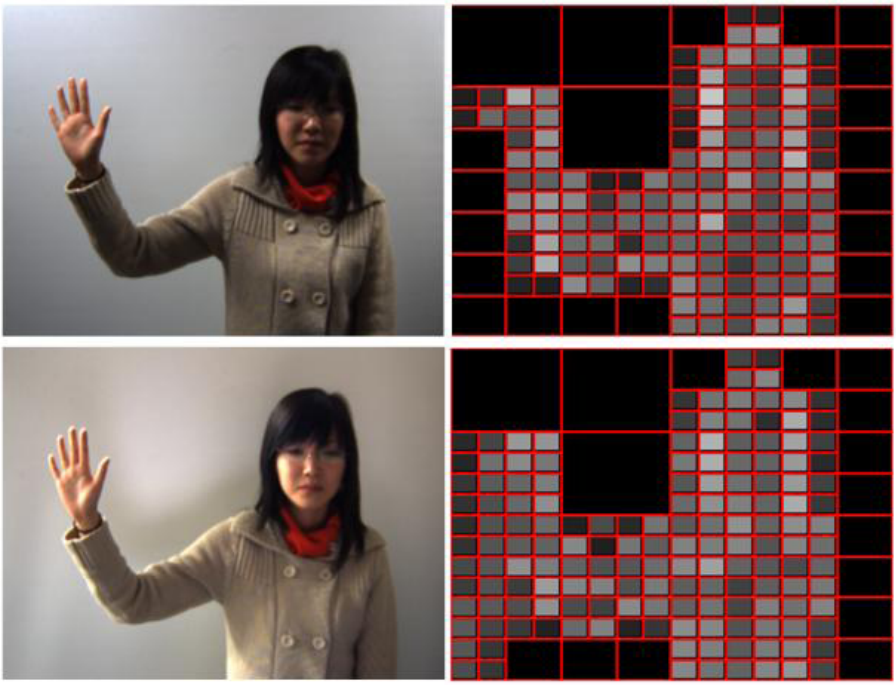


Fig. 4. Foreground cell extraction using multi-scaled search algorithm of test image with light on/off. The maximum search level is 4, the level threshold *Tl* is 0.03.

3 Vision-Game Interface

The extracted foreground cells give enough features to be recognized as a distinctive pose to be transferred to game system as an input event. We consider gaming environment that the user sits close to the monitor and the camera, then the absolute position of foreground cells from the user's upper body are used as inputs for the simple experimentation. If the captured signal falls in a predetermined set, the recognition module compares the current foreground cells with the each of the stored templates and either selects the category of the closest match or interpolates between templates using Euclidian distance or similar measurement. Another obvious method is to collect an

example database of poses and use it to train an artificial neural network. This avoids having to define each gesture in detail. It is hoped that the network will find the important features and abstract them. In addition, the networks are often robust to noise and work well for different individuals.

3.1 Recognition with Neural Network

Local histogram approach can be considered as a natural reduced feature vector representing the coarse pose on the screen. Therefore the v×h foreground probability cells can be easily applied to the inputs of neural network to classify an output pattern rather than using whole pixels.

With the variation to the position of both arms of the user, we assign the 7 target poses as shown in Figure 5. As described in next sub-section we place a number of grouped poses for different input map. The neural network is structured with the v×h number of inputs from the foreground probability cells, 1 hidden layer consisting of 3 or 4 neurons and r classified outputs according to the required output number. By grouping required poses into a group, the network can give more discriminative power. We train the neural network with the 100 training data set for each pose using the back propagation method. The training data sets are captured from 2 adult persons, each a man and a woman.

3.2 Interfacing to Game Engine

The whole previous works must be integrated to work with existing game engine. To communicate with the existing game, the seamless connection is preferred. Since most PC game requires the keyboard and the mouse as input devices, it would be perfect if the classified output from the neural network is translated into keyboard or mouse input event. For higher integration at application level, the game engine should prepare a number of different input map that describe a number of input events and corresponding event handlers needed for current game context. Also, the interfacing library limits the maximum number of recognizable poses so that the game input map can be adequately prepared. Each input map represents the mapping between the input event and its handler under different game context. For example, the enter key of keyboard, which come from first pose of input map #1, can trigger the selecting the start menu in main menu screen or the firing a missile in game playing scene. Or, for the same game context, the different poses can be set for user convenience and it will help gamers immerse into virtual environment. Given that, the game engine should notify the current game context to the interface system to load different training set or initialize system according to the required number of input events.

We separate the vision-related modules from interfacing module for performance reason. Although the integral histogram and multi-scaled search algorithm boost the overall computation speed, the low-powered CPU show poor frames per second in 3D game if the recognition process run at the same system. So the vision processing module can be launched on the independent system and send the result to the game via network socket connection, the gamer can play high-end game without loss of performance.

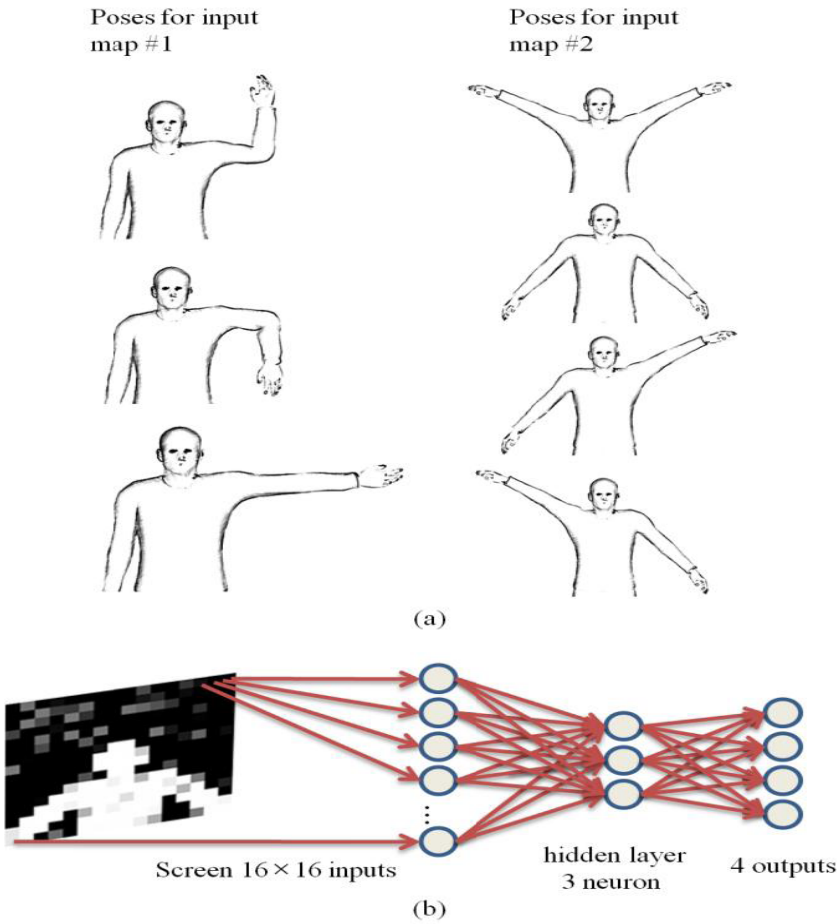


Fig. 5. (a) Selected poses. (b) Neural network architecture for input map #2.

4 Experiments

We have built the vision game interface library written in C++. The main idea of this paper is implemented with OpenCV vision library. We have developed a simple game using Torque3D to show the usability of the proposed system in the real-time processing situation. The game application is free flying simulation over a simple terrain map with first person view. The camera view position drifts freely until the neural outputs are produced. Before the game starts, the simple main menu comes out for selecting different map where the system use different input map from game.

playing scene. In the game playing scene, when foreground objects detected, the neural classifier produces one of predefined classes which converted to keyboard input and the result is transmitted to the game system. With the classified pose pattern, the user can turn left/right or accel/stop in the game playing scene.

The camera grabs 320×240 RGB images at 60fps and the testing PCs are equipped with Intel dualcore 2.0G processor, 1G memory and Radeon X1600 graphic card for both the game and recognition processes. We measured 30 fps with 320×240 sized integral histogram process. Moreover, with resized 160×120 image, we get 60 fps, i.e. sufficient for application to 3D gaming environments. The main computational load is scanning process of integral histogram where we expect that SIMD processing can help significant performance improvement.

Table 1. Recognition results for each pose

Pose #	Correct	Error	Detection Ratio
1	20	1	95
2	25	2	92
3	15	1	93
4	53	4	93
5	48	3	94
6	64	7	90
7	55	5	91
Total	280	23	92

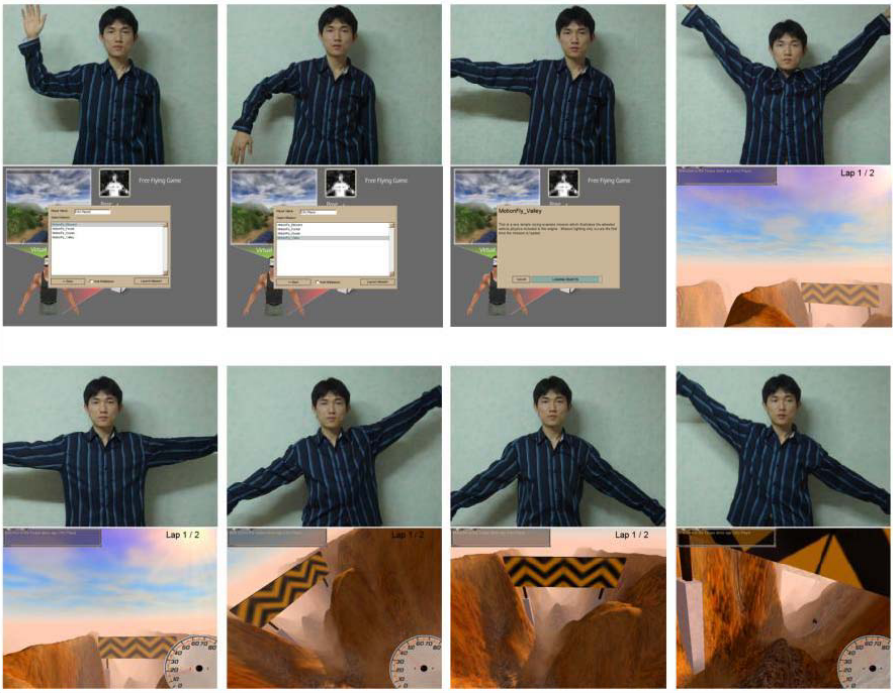


Fig. 6. Tracking results sampled every 90 frames

We prepared the ground truth data from captured movie clip. The 560 frames of image are captured and the expected pose is manually marked at each frame. We compared the ground truth data to recognition result and Table 1 shows the recognition result. The poses from 1 to 3 are used for menu traveling in the menu screen and others are used for controlling the air-plane view. Note that the frames that have non-determined poses are excluded from counting. The overall detection ratio is more than 90% so that the proposed vision-based interface system enables the game user to actually control the game without difficulty. Figure 6 shows the both screens from the camera and the game scene.

5 Conclusions

In this paper we extract foreground cells using local oriental histogram comparison and the extracted information is used as inputs to the trained neural network for pose recognition. For robust and fast computation of local histogram we perform Gaussian kernel to 1D orientation histogram and utilize the computational efficiency of integral histogram. In addition, multi-scaled search algorithm proved to be tolerant to camera noise and slight moving object including illumination changes in background. The effective way of processing input events from the vision based module is proposed and implemented as a vision interface library and the simple 3D game is given as a test bed to prove its efficiency.

Acknowledgments

This research was supported by the ITRC (Information Technology Research Center, MIC) program and Seoul R&BD program, Korea.

References

1. Jang, D.H., Chai, Y.J., Jin, X.H., Kim, T.Y.: Realtime Coarse Pose Recognition using a Local Integral Histogram. In: International Conference on Convergence Information Technology, November 21-23, pp. 1982–1987 (2007)
2. Mason, M., Duric, Z.: Using histograms to detect and track objects in color video. In: Applied Imagery Pattern Recognition Workshop, AIPR 2001 30th, IEEE, October 10-12, pp. 154–159 (2001)
3. Porkili, F.: Integral histogram: A fast way to extract histograms in cartesian spaces. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (2005)
4. Noriega, P., Bascle, B., Bernier, O.: Local kernel color histograms for background subtraction. VISAPP, 1st edn., pp. 213–219. INSTICC Press (2006)
5. Noriega, P., Bernier, O.: Real Time Illumination Invariant Background Sub-traction Using Local Kernel Histograms. In: British Machine Vision Association (BMVC) (2006)
6. Bradski, G.: Real time face and object tracking as a component of a perceptual user interface. In: Proc. IEEE WACV, pp. 214–219 (1998)
7. Viola, P., Jones, M.: Robust real time object detection. In: IEEE ICCV Workshop on Statistical and Computational Theories of Vision (2001)

8. Stauer, Grimson, W.E.L.: Adaptive background mixture models for real-time tracking. In: Computer Vision and Pattern Recognition Fort Collins, Colorado, June 1999, pp. 246–225 (1999)
9. Wixson, L.: Detecting salient motion by accumulating directionary-consistent flow. IEEE Trans. Pattern Anal. Machine Intell. 22, 774–780 (2000)
10. Yahav, G., Iddan, G.J., Mandelbaum, D.: 3D imaging Camera for Gaming Application. In: Consumer Electronics 2007. ICCE 2007. Digest of Technical Papers (2007)
11. Marimon, D., Ebrahimi, T.: Orientation histogram-based matching for region tracking. In: IEEE Eight International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2007) (2007)
12. Viola, P., Jones, M.: Robust real time object detection. In: IEEE ICCV Workshop on Statistical and Computational Theories of Vision (2001)
13. Li, L., Huang, W.M., Gu, I.Y.H., Tian, Q.: Foreground object detection in changing background based on color co-occurrence statistics. In: Proc. IEEE Workshop Applications of Computer Vision, December 2002, pp. 269–274 (2002)
14. Freeman, W., Tanaka, K., Ohta, J., Kyuma, K.: Computer vision for computer games. In: Int'l Workshop on Automatic Face- and Gesture-Recognition (1996)

Memotice Board: A Notice Board with Spatio-temporal Memory

Jesús Ibáñez[1], Oscar Serrano[1], David García[1], and Carlos Delgado-Mata[2]

[1] Departamento de Tecnología, Universidad Pompeu Fabra, Barcelona, Spain
{jesus.ibanez, oscar.serrano, david.garcian}@upf.edu
[2] Universidad Panamericana, Aguascalientes, CP 20290, México
cdelgado@ags.up.mx

Abstract. This paper describes the design and development of a novel digital notice board which allows non experienced users to easily interact with digital information. In particular, the system allows the user to receive and handle media elements (pictures, text messages, videos). Instead of employing the file system to interact with information, the user interface promotes a kind of interaction which relies on spatial and temporal memory, which we believe to be more adequate.

1 Motivation

The area of technology-enhanced learning deals, normally, with the tasks and objects more directly involved in the learning process (contents, exercises, evaluation, etc). However, in a conventional physical education centre, there are other objects and tasks which, although not directly related to the learning process, enhance the student experience. An example is a notice board in an education centre which represents a social space where students' interests meet. The notice board is not only useful because of its main functionality, but also because of its social side. We research new approaches to include this kind of objects and tasks in the technology-enhanced learning process in an appropriate way, as we think that these objects can also benefit from novel technology-enhanced approaches.

In particular, in this paper we describe the design and development of Memotice Board, a digital notice board with spatio-temporal memory. The design of this system is based on previous work we carried out in the framework of the EC funded IST project ICING (Innovative CIties of the Next Generation) which, among other things, explores new ways of communication and interaction. In this context, needs for communication mechanisms and social awareness were found in several communities (elderly people, women association, families, etc.) through user studies. In order to fulfill these requirements we designed DINDOW (DIgital wiNDOW), a system which allows the user to receive, handle and send media elements (pictures, text messages, videos) in a very simple way. Instead of employing the file system to interact with information, the user interface promotes a kind of interaction which relies on spatial and temporal memory, which we believe to be more adequate for our users. Thus, DINDOW is our original

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 401–409, 2008.

and generic application for communication and social awareness in communities. We later extended and adapted DINDOW to be employed in education centres. The result was Memotice Board, the system which is described in this paper, that among other things adds the possibility of defining different levels of information access.

The next subsections begin with a review of related work. Following this we describe the design and development of the system, including its architecture, user interface and interaction design. Finally we provide the conclusions and future work.

2 Related Work

Awareness systems, according to Markopoulos et al. [10], can be defined as systems whose purpose is to help connected individuals or groups to maintain a peripheral awareness of the activities and situation of each other. The area of awareness system is a flourishing field of research, and interesting systems have been proposed, in the last years, for both workplaces [6][9] and social/family life [8][11][12][5]. In this sense, Memotice Board is an awareness system which on the one hand can be used as a peripheral display (it can be used as a notice board displaying pictures, news, etc) and on the other hand can be actively employed (the students can interact with the interface and the system administrator can manage the collection of multimedia elements).

It is well known that managing disparate data through traditional hierarchical storage and access interfaces is frustrating for users [3], specially for non experienced ones. As a consequence, different approaches and metaphors have been proposed to replace the desktop metaphor and its related hierarchical file system. A specially interesting alternative was proposed by Fertig et al. in Lifestreams [7], which uses a timeline as the major organizational metaphor for managing file spaces. It provides a single time-oriented stream of electronic information, and supports searching, filtering and summarization. Rekimoto extended this uni-dimensional idea to a bi-dimensional interface in TimeScape [14], which combines the spatial information management of the desktop metaphor and time travelling. Our system, Memotice Board, extends these ideas by allowing time-travelling in any subregion of the user interface, and adding the ability to formulate natural and intuitive spatio-temporal queries.

3 Design and Development

In this section we describe the design and development of Memotice Board. We first outline its overall architecture. Following this, we present the user interface and the interaction of the system.

3.1 Architecture

The overall architecture of the system is shown in figure 1. It is composed of two kinds of agents (administration agent and public agent), a database and a file system.

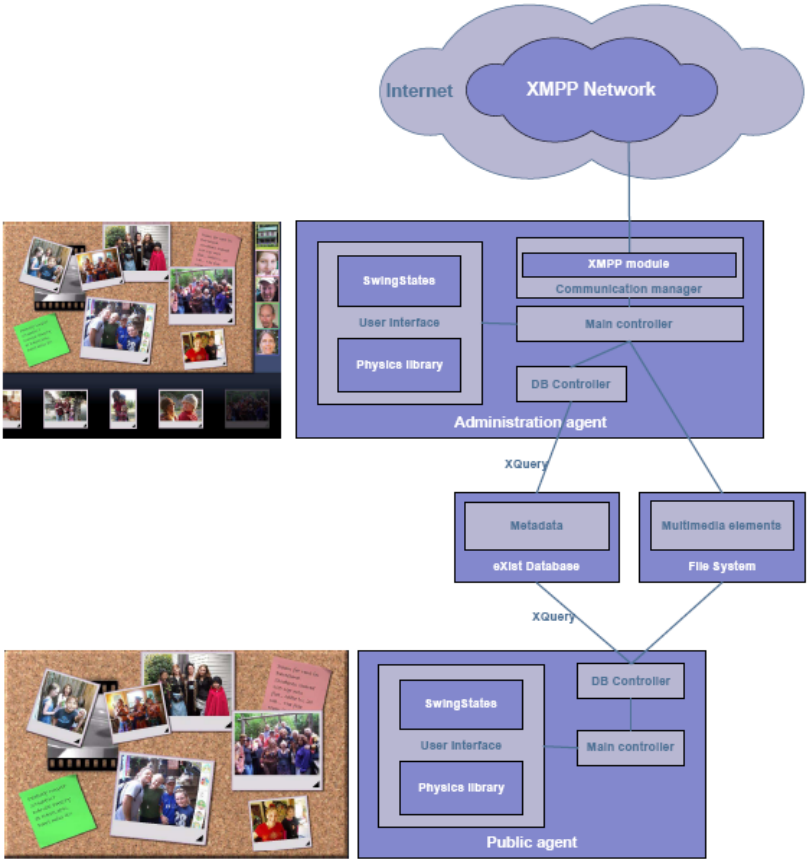


Fig. 1. Overall architecture

The administration agent is the main agent, which is employed by the administrator to manage the complete system. This agent includes the communication manager that is responsible for receiving and sending information. It consists of a series of modules implementing di_erent communication modalities. The current version contains a module implementing communication over the XMPP protocol. Thus, users from the Internet (students, ex-students, teachers, etc) can send new elements (text messages, images or videos) to the Memotice Board by using an XMPP client (that is, an instant messenger client). By using the administration agent, the administrator can access all the functionalities of the system.

The information stored by the system can be classified into two categories: multimedia elements and metadata. The multimedia elements are the images and videos received by the system from other users. These elements are automatically stored in the file system. The user does not need to know where these elements are physically located nor how they are named. The metadata are semantic annotations about the multimedia elements. These annotations include, for instance, the date when the element was received, its width and height, its location on the screen, its location on the file

system, etc. The metadata are stored in a database. In particular, in the current version the metadata are stored in eXist [13], an open source native XML database. Thus, this metadata is accessed through queries programmed in the XQuery language [16].

There can be any number of public agents. Each public agent is responsible for controlling a particular digital interactive notice board which is exposed in a public space in the education centre so the students can interact with it. Usually, a public agent allows the user only to access some of the system functionalities (a subset of the functionalities provided by the administration agent).

The user interface (of both kinds of agents) incorporates novel interaction mechanisms, programmed with SwingStates. The user interface also employs a physics library to simulate the behaviour of group movement. The combination of both the novel interaction mechanisms implemented in the user interface and the metadata stored in the database, allows the user to naturally and intuitively interact with the multimedia elements.

3.2 User Interface and Interaction

As stated by Beaudouin-Lafon [2], the only way to significantly improve user interfaces is to shift the research focus from designing interfaces to designing interaction. This requires, among other things, powerful interaction models, beyond the usual ones. Moreover, as we were prototyping following an iterative process to design the system, we required advanced software libraries which ease the rapid development of new kinds of interaction. The typical libraries (like Java Swing) based on a set of well known graphical widgets are not appropriate, as they are oriented to create new user interfaces, not new interaction mechanisms. After surveying the state of the art, we finally decided to employ SwingStates [1]. SwingStates is a library that adds state machines to the Java Swing user interface toolkit. Unlike traditional approaches, which use callbacks or listeners to define interaction, state machines provide a powerful control structure and localize all of the interaction code in one place.

The user interface of the public agent is a subset of the user interface of the administration agent. Thus, we will start by describing the latter one, and we will later specify which parts of it compose the former one. Figure 2 shows the user interface of the administration agent. As shown in the figure, the user interface is composed of three main regions (A, B and C). Region A is a scrollable space where the multimedia elements (pictures, videos, texts) are automatically added when they are received by the system. In particular, when a new element arrives, it appears on the right of this area, while the other elements in this space shift to the left accordingly. Thus, all the elements in region A are temporally ordered. An element X was received earlier than an element Y if X is visually located to the left of Y . The user can scroll through the elements of this space. Thus, he can access any received element.

When a new element is received, not only does the element appear on the right of region A, but also is the user warned by a particular sound. Moreover, the new element is enhanced with a colourful frame which identifies it as a new incoming element.

The user can group elements in folders but always keeping the temporal order. Thus, the folders are created in a simple way. The user only decides the extreme elements of the folder, that is, the oldest and newest element, and the folder is automatically created.

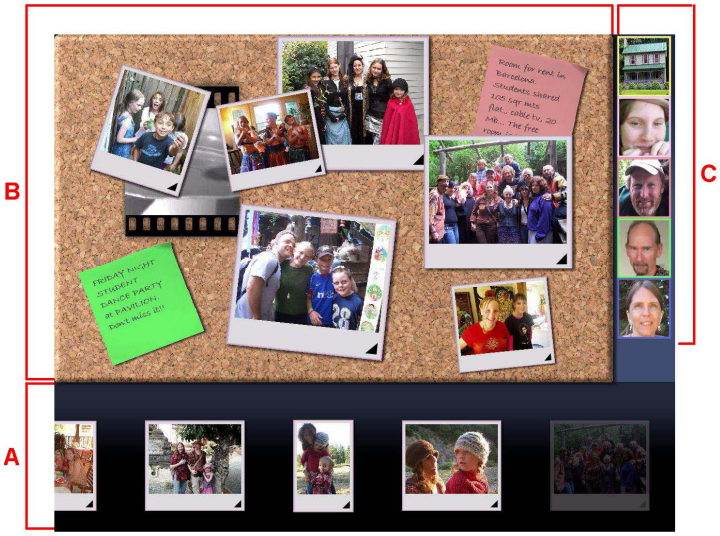


Fig. 2. Administrator user interface

Fig. 3. User interface with closed and open folders

Folders also appears temporally ordered on region A. As shown in figure 3, a closed folder is represented as a photo album such that its front cover is illustrated with its most representative content (the element which has been publicly exposed for the longest time). When a folder is open, its contents are highlighted with the folder colour.

When several elements arrive together in the same message, they are automatically associated as a group. Then, these elements will behave as a group. For instance, the user can move the whole group by just moving one of the elements. The smooth movement of the group is simulated by employing the proper algorithms from a physics programming library.

Note that pictures, videos and text messages are clearly differentiated by the way they are represented. Pictures are displayed with the Polaroid aesthetic, while videos are presented with a film look, and text messages are shown in coloured PostIts such that the colour identifies the kind of sender (that is, there is a particular colour per each sender: student, ex-student, professor, education centre, etc).

Region B is an area which can be seen as a kind of advanced notice board. The user can move multimedia elements from region A to this area. When that happens, the elements which are moved to region B still remain visible in region A, although in ghost mode (with a degree of transparency). Elements moved from A to B, appear bigger in B than in A. However, relative size among elements is respected. That is, an element X will appear bigger than an element Y in region B if X as well appears bigger in region A.

The elements in region B will keep visible until the user decides to remove them. In this region, the multimedia elements can be freely moved, rotated and scaled. Moreover, video elements can be played and text annotation can be added to pictures.

Region B has a kind of spatio-temporal memory. For each element, its state-changes (location, size, rotation angle, time of change,...) are annotated in the database. Thus, by employing this memory it is possible to time-travel in this region. By travelling to the past, this region evolves by showing previous states of the region in an inverse time order. Elements appear and disappear, at a configurable rate, at the locations they were located at previous times, giving the user the impression that the region is travelling to past. Moreover, it is possible to time-travel in a subregion of region B. The user can mark a particular subregion by drawing its border on the screen. As a result, a hole appears in the just marked subregion. Then the user can time travel only in this subregion (see figure 4). Thus, he can access, for instance, an ad about a flat for renting he remembers that was located on the center of the right side by marking this area and time-travelling to past. Time-travels to the future (from past times) are also allowed. Both, the direction of the time-travel (forward or backward) and its speed are controllable through a visual widget.

Natural spatial queries can be formulated in region B. To start the query, the user marks a subregion by drawing it on the screen. Then he can specify search criteria in an intuitive way. For instance, he can specify the sender of the element he is searching by dragging the picture of the sender from region C to the just marked subregion. He can also specify the approximate size of the element being searched by just drawing it. Thus, queries are constructed by using simple gestures. In the mentioned example, the system will find and show elements which were located in the marked subregion and fulfill the specified search criteria (sender and approximate size).

Region C is a space that contains a set of pictures representing the different kinds of contacts (students, ex-students, teachers, education centre, etc). These pictures can be used as clues to filter (for instance to look for elements sent by a particular type of contact).

Fig. 4. Time-travel in a subregion

As mentioned before, the user interface of the public agent is a subset of the user interface of the administration agent. The administrator is responsible for defining the capabilities of each public user interface. Usually, it consists only of the region B, that is, the proper notice board (see figure 5). Thus, students can interact with the published elements by time-travelling and performing spatial queries. However, they cannot remove any element from the notice board (this is the administrator's responsibility).

Fig. 5. Typical public user interface

4 Conclusions and Future Work

This paper has presented Memotice Board, a novel digital notice board which allows non experienced users to easily interact with digital information. Instead of employing the file system to interact with information, Memotice Board promotes a kind of interaction which relies on spatial and temporal memory, which we believe to be more adequate for our users.

Future work includes the evaluation of the system as an awareness system. This evaluation will be based on both the ABC (Affective Benefits in Communication) questionnaire [15] and the IPO-SPQ (IPO Social Presence Questionnaire) [4].

Acknowledgments

The authors thank both Leticia Lipp for generously proofreading and Caroline Appert for her quick responses to our doubts when using SwingStates. This work has been partially funded by the European Union IST program through the project "ICING: Intelligent Cities for the Next Generation".

References

[1] Appert, C., Beaudouin-Lafon, M.: Swingstates: adding state machines to the swing toolkit. In: UIST 2006: Proceedings of the 19th annual ACM symposium on User interface software and technology, pp. 319–322. ACM Press, New York (2006)

[2] Beaudouin-Lafon, M.: Designing interaction, not interfaces. In: AVI 2004: Proceedings of the working conference on Advanced visual interfaces, pp. 15–22. ACM Press, New York (2004)

[3] Cook, T.: It's 10 o'clock: do you know where your data are? Technology Review 98(1), 48–53 (1995)

[4] de Greef, P., IJsselsteijn, W.A.: Social presence in a home tele-application. CyberPsychology and Behavior 4, 307–315 (2001)

[5] Dey, A.K., de Guzman, E.: From awareness to connectedness: the design and deployment of presence displays. In: CHI 2006: Proceedings of the SIGCHI conference on Human Factors in computing systems, pp. 899–908. ACM Press, New York (2006)

[6] Dourish, P., Bly, S.: Portholes: supporting awareness in a distributed work group. In: CHI 1992: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 541–547. ACM Press, New York (1992)

[7] Fertig, S., Freeman, E., Gelernter, D.: Lifestreams: an alternative to the desktop metaphor. In: CHI 1996: Conference companion on Human factors in computing systems, pp. 410–411. ACM Press, New York (1996)

[8] Hindus, D., Mainwaring, S.D., Leduc, N., Hagstrm, A.E., Bayley, O.: Casablanca: designing social communication devices for the home. In: CHI 2001: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 325–332. ACM Press, New York (2001)

[9] Jancke, G., Grudin, J., Gupta, A.: Presenting to local and remote audiences: design and use of the telep system. In: CHI 2000: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 384–391. ACM Press, New York (2000)

[10] Markopoulos, P., de Ruyter, B., Mackay, W.E.: Awareness systems: known results, theory, concepts and future challenges. In: CHI 2005: CHI 2005 extended abstracts on Human factors in computing systems, pp. 2128–2129. ACM Press, New York (2005)
[11] Markopoulos, P., IJsselsteijn, W., Huijnen, C., Romijn, O., Philopoulos, A.: Supporting social presence through asynchronous awareness systems. In: Being There -Concepts, Effects and Measurements of User Presence in Synthetic Environments, pp. 261–278. IOS Press, Amsterdam (2003)
[12] Markopoulos, P., Romero, N., van Baren, J., IJsselsteijn, W., de Ruyter, B., Farshchian, B.: Keeping in touch with the family: home and away with the astra awareness system. In: CHI 2004: CHI 2004 extended abstracts on Human factors in computing systems, pp. 1351–1354. ACM Press, New York (2004)
[13] Meier, W.: eXist: An Open Source Native XML Database. In: Chaudhri, A.B., Jeckle, M., Rahm, E., Unland, R. (eds.) NODe-WS 2002. LNCS, vol. 2593, pp. 169–183. Springer, Heidelberg (2003)
[14] Rekimoto, J.: Time-machine computing: A time-centric approach for the information environment. In: ACM Symposium on User Interface Software and Technology, pp. 45–54 (1999)
[15] van Baren, J., IJsselsteijn, W.A., Romero, N., Markopoulos, P., de Ruyter, B.: Affective benefits in communication: The development and field-testing of a new questionnaire measure. In: PRESENCE 2003, 6th Annual International Workshop on Presence, Aalborg, Denmark (October 2003)
[16] Walmsley, P.: XQuery. O'Reilly Media, Inc., Sebastopol (2007)

Mobile Cultural Heritage: The Case Study of Locri

Giuseppe Cutrí[1], Giuseppe Naccarato[2], and Eleonora Pantano[3]

[1] University of Turin, 10124, Turin, Italy
giuseppe.cutri@unito.it
[2] University of Calabria, 87036, Arcavacata di Rende, Italy
naccaratogiuseppe@gmail.com
[3] University of Calabria, 87036, Arcavacata di Rende, Italy
eleonora.pantano@unical.it

Abstract. The goal of this project is to study the use of mobile technologies equipped with global positioning systems as an information aid for archaeological visits. In this study we will focus in the study of the technologies used to implement these systems. To this end we analyze an archaeological site where this systems have been tested. In this experiment we have applied state of the art technologies in virtual and augmented reality to implement a system that allows users to access the site using their mobile devices. We conclude that the use of this kind of technologies is an effective tool to promote the archeo-geographical value of the site.

Keywords: mobile device, mobile virtual navigation, digital reconstruction, GPS, cultural heritage.

1 Introduction

Advances in mobile technologies are enjoyed by an increasing percentage of the population. This is due mainly to lower prices and to the technologization of life and work style standards of the population [1].

Most of the current communication processes are based on the use of mobile devices. Some of the most used are tablet pc, pocket pc, smart-phone, PDA (Personal Digital Assistant), and iPod. These technologies provide several web tools like, search engines, virtual communities and e-advertising among others.

Adapting the power of these technologies to the field of cultural heritage, allows the broadcast of local heritage to a worldwide level. Innovative uses of technology can stimulate curiosity and interest in users, satisfy their information needs and ultimately allow the creation of a digital heritage [2] [3]. These devices can guide users in virtual or real world spaces. Virtually reconstructed environments take advantage of information rich databases providing the users with historical, cultural, and geographical data. In these environments the user can better explore in an augmented reality space. This system empowers the user giving a knowledge rich environment that facilitates learning [4] [5] [6].

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 410–420, 2008.

This type of systems can mix real and virtual worlds, allowing the combination of the geographical location with the exact historical or cultural information. An added tool that identifies the geographical position [8] allows the system to have a combined view of a culturally interesting artefact with its virtual reconstruction (3D model). This 3D model represents its original shape. The user can also benefit by using the user friendly interface of the device, to view other multimedia data related to the artefact, for example the reconstruction of the virtual reproduction of the original environment and the historical source [9].

The aim of this paper it to highlight the possibility to apply these technologies to regions such as Calabria, which is rich in cultural and archaeological resources not always exploited. In particular we study the system at the archaeological park of Locri and study how the tools we describe can improve and value the enjoyment of the place.

2 Mobile Virtual Navigation

New mobile devices are becoming more and more popular due to their low cost and their advantages in connection to new services and social interaction. These devices are not mere cell phones or organizers, but powerful computing devices [7]. In this study we have used a mobile implementation of a Virtual Navigation System (VNS). Current VNSs were developed as a desktop application to simulate visits for a wide range of environments, ranging from a reconstructed city to a museum. Today the high performance of mobile devices offers the possibility to combine the capabilities of a desktop Virtual Navigation System with the ones of a Global Positioning System (GPS) device. Using a mobile device with GPS (often already integrated on most PDAs and cell phones) the VNS provides more exciting features such as allowing the user to have real and virtual information combined depending on its location. This system has been developed with the goal of offer a better experience while visiting archaeological sites.

3 System Architecture

The system we present is a program that enriches the exploration of open spaces with additional data. The system provides real time visualization, on a mobile device, of a 3D reconstruction of the environment. This environment also gives navigation capabilities using its GPS data system.

To achieve this task we designed a new graphical engine for the mobile device. The system is built on top of a set of new graphic libraries developed in collaboration with the E-Guide S.R.L. which were called Q3 libraries. These C++ APIs (Application Program Interfaces) are divided in two parts:

- Q3Engine: is a 3D graphic engine;
- Q3Widgets: is a GUI library for rendering the GUI (Graphic User Interface).

These two set of APIs are linked together by these other libraries:

- Q3Toolkit: is the glue between the 2D layer and the 3D layer and the OS (Operating System);
- Q3Lib: offers many platform independent functions as well as computational geometry functions used to manipulate meshes and other 3D and 2D objects;
- Q3GPS: receive and process GPS data.

The API is built on top of OpenGL ES and OpenVG [10] libraries that are the standard de facto in mobile environments (Many CPU manufacturers support natively these libraries on their products). A converter allows the import of a COLLADA [11] file or a Google Earth file (Google Earth 4 files are compressed COLLADA files, with textures and other information) and save them as a compressed format specifically designed for.

The system processes GPS data to obtain the user position and move the virtual environment along with the user's movements. It is also possible to connect a GPS with an integrated compass in order to know the user's orientation. If the compass data is missing the user has the possibility to move the view using the joypad of the device. The meshes position is stored in xml format together with other information such as a text and multimedia contents so the user can click on any object of the world and read the description, watch images, and so on.

Since most of today's mobile devices do not have FPU (Floating-Points Unit) the system hasn't the possibility to be compiled using fixed-points and a special library for fixed-point algebra was developed.

The GUI has many graphical effects like shading, anti-aliasing and is planned to be used using a touch screen. The 3D engine can show any textured mesh and support multiple light effects.

4 The Case Study of Locri

We chosed to analyze the case study of Locri because of the rich archaeological heritage of the zone and because, at the moment, it hasn't been studied yet, using the latest technologies.

Locri Epizefiri is one of the most important Greek poleis of Calabria. Its archaeological park covers a big area: more than 568,34 acres. It spreads out along the coast and the mountains (Fig. 1 shows the findings of famous Greek Theatre built the IV century b.C.)[12], [13].

To implement the systems for the archaeological site we followed several steps:

- Evaluation of findings accessibility;
- Evaluation of most interesting routes inside the park;
- Access to useful information to reconstruct ancient artefacts.

Fig. 1. An image from the archaeological site of Site

Fig. 2. The zone of Centocamere, in the archaeological site of Locri (image from Google Earth)

Inside the archaeological park of Locri, tourist can visit 3 zones: Centocamere (Fig. 2), Museum (Fig. 3), Theatre (Fig. 4).

In these zones there are few routes that allow tourists to access the most interesting findings. Archaeologists have excavated the ancient ruins especially in the zone of Centocamere, where the ruins of the ancient city centre are located. These are characterized by houses, and workshops where clay ceramics were manufactured and sold.

For example, in Fig. 5 we show the map of the zone of Centocamere in the park, where we highlight the possible routes:

Fig. 3. Museum zone, in the archaeological site of Locri (image from Google Earth)

Fig. 4. Theatre zone, in the archaeological site of Locri (image from Google Earth)

We investigated the required details to develop the virtual reconstructions of the ancient objects of the zone to test and validate the system. This test shows several routes that a tourist can visit (Fig. 6). The system gives the user also the opportunity to choose a fixed route from the list of all possible ones. In fact, the user can visit the archaeological park with his personal mobile device and use it to choose his preferred route to visit the park. For example the user can choose a fixed route or can invent his own choosing the most interesting findings to see.

The most important part of the visit is the route and it is a fundamental factor to exploit the territory. From a mathematical point of view, we can describe the

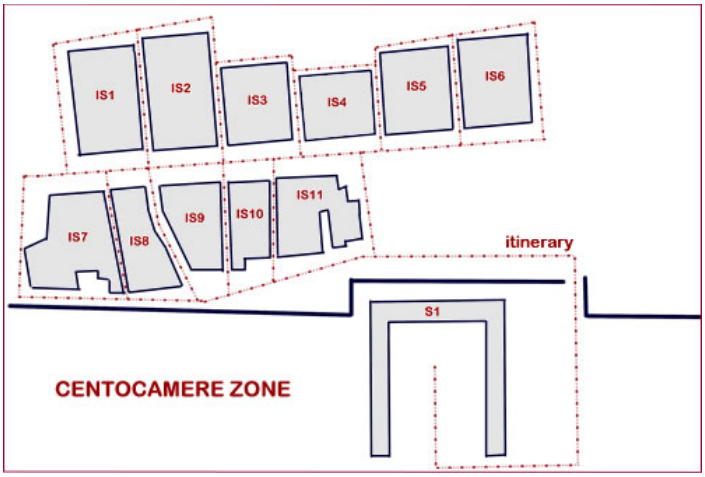


Fig. 5. A map with the possible routes of zone of Centocamere in the archaeological park of Locri

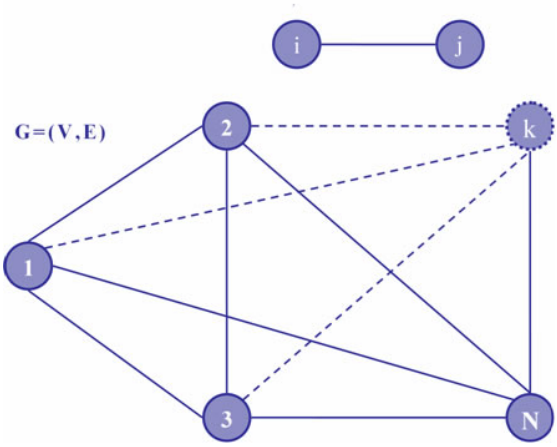


Fig. 6. Mathematical representation of place of interest and connections among them

place by using a graph G, defined G=(V,E), where V is a set of vertices and represents the places of interest and E is the set of links and represent the possible connections among them (the communication channels) [14].

After the mathematical formalization of the routes, we can apply mathematical tools to find the itinerary that maximizes the travel performance and a more personalized route [14].

It is possible to use technologies of virtual reality and computer graphics to use in terrains in an efficient way. These allow to reconstruct archaeological sites and environments which existed only in the past [15]. The traditional access to archaeological ruins required a mental effort from visitors because they had to reconstruct in their minds the ancient scenario. Using this system the virtual

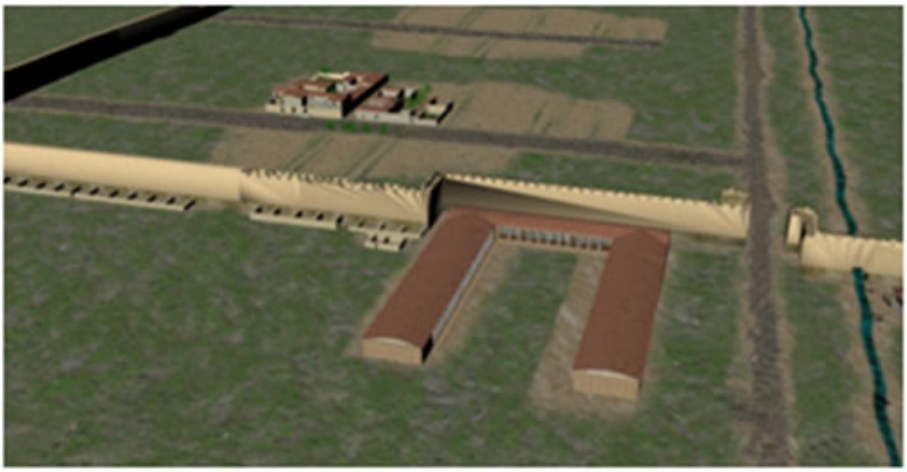


Fig. 7. Virtual reconstruction of the Centocamere zone

Fig. 8. An example of the user-friendly interface for language choice

reconstruction of objects and environment using graphics, audio/video reproduction allow users to live a more interesting and immersive experience [16] [18]. Virtual reconstructions and their related multimedia contents make visit more interesting and instructive (Fig. 7).

4.1 How the System Is Working

In may museums or archaeological sites, tourists can find audio-guides which guide them along fixed routes, or force them to use information points with a pc in which visitors can get access interactive information. In this paper we present

Fig. 9. An example of the user friendly interface to choose the route

Fig. 10. Stoá in the zone of Centocamere

a different tool, because it is not stationary, it can be personalized by user, and it is based on the geographical position of the user.

Mobile devices, and wireless communication systems, are combined with virtual and augmented reality to obtain a new tool which can be an electronic, personalized and mobile guide through archaeological sites [17].

We can summarize the use of this system in the following fundamental steps:

- STEP 1: User accesses to archaeological site and decides to rent the particular mobile device or to use his own (in this case he has to download on his device all the useful information, like maps, photos or other data);

Fig. 11. Access to Stoá from mobile device

- STEP 2. The user starts the application and chooses the language (Fig. 8) and route (a fixed route from the list or personalize his own) (Fig. 9);
- STEP 3. The device becomes a tourist guide. It locates the geographical position of user (using a GPS system) in the park. When the user is close to a particular object (Fig. 10), the display shows a virtual reconstruction. The user can see the real object while comparing with the reconstruction in the mobile device(Fig. 11). The user can play the object (as a game) and he can choose to listen to historical data, or information about the structure or manufacturing process, read the text or visualize other multimedia information.

5 Conclusion

We illustrate how the archaeological park of Locri, and is big extension, can be enjoyed in a more effective and efficient way by using this new system. This system allows users to understand, learn and appreciate also parts that don't exist anymore, artefacts, which were destroyed by weather or man. The user can experience an immersive and more interesting experience, especially for that part of population which is less interested in the archaeological heritage but more sensitive to the use of new technologies [19].

A similar system could be applied to other archaeological sites with the same success.

6 Future Work

Next generation of mobile devices will have more powerful CPUs and many will have a GPU also, which means that there will be no problem rendering very

complex meshes with a low frame rate. These new kind of devices will allow to render even more realistic scenes.

But the future is not only based on new powerful hardware. AR (Augmented Reality) systems will play an important rule on the current scenario. Even today a lot of mobile phones and PDAs have a camera inside which already permits to embed real image data in order to overlap the 3D reconstruction on top of the reality. People can experience new HCI (Human Computer Interaction) that will permit a more interaction with the environment and many new exciting features.

The next step will be the integration of maps and a multi-modal guidance engine which will permits the user to be guided throw a city using various transport services (bus, train, taxi, etc.). These technologies are currently under development in collaboration with the E-Guide S.R.L.

Furthermore, the system we presented will be tested and evaluated through a quantitative analysis with consumers in the archaeological park of Locri.

References

1. Reitano, A., Pantano, E., Feraco, A.: Comunicazione digitale e gestione del territorio (in press, 2007)
2. Parry, R.: Digital heritage and the rise of theory in museum computing. Museum management and Curatorship 20, 333–348 (2005)
3. Lin, Y., Xu, C., Pan, Z., Pan, Y.: Semantic modeling for ancient architecture of digital heritage. Computers & Graphics 30, 800–814 (2006)
4. Bilotta, E., Pantano, P., Rinaudo, S., Servidio, R.C., Talarico, A.: Use of a 3D Graphical User Interface in Microelectronics Learning and Simulation of an Industrial Application. In: Proc. 5th Eurographics Italian Chater Conference, pp. 217–224 (2007)
5. Pan, Z., Cheok, A.D., Yang, H., Zhu, J., Shi, J.: Virtual reality and mixed reality for virtual learning environments. Computers & Graphics 30, 20–28 (2006)
6. Cai, Y., Lu, B., Zheng, J., Li, L.: Immersive protein gaming for bio edutainment. Simulation & gaming 37(4), 466–475 (2006)
7. Bellotti, F., Berta, R., De Gloria, A., Margarine, M.: MADE: developing edutainment applications on mobile computers. Computer & Graphics 27, 617–634 (2003)
8. Burrogh, P.A.: Principles of geographical information systems for land resource assessment. Clarendon Press, Oxford (1986)
9. Gleue, T., Dhne, P.: Design and Implementation of a Mobile Device for Outdoor Augmented Reality in the ARCHEOGUIDE Project. Virtual Reality. In: Archaeology, and Cultural Heritage International Symposium (VAST 2001) (2001)
10. OpenGL ES and OpevVG (2007), http://www.khronos.org/
11. COLLADA (2007), http://www.khronos.org/
12. Costamagna, L., Sabbione, C.: Una cittá in Magna Grecia Locri Epizefiri. Guida Archeologica. Laruffa Editore. Reggio Calabria (1990)
13. Serafino, C.: Locri antica e il suo territorio. Il Portichetto. Aga- Cuneo (1991)
14. Bertacchini, P.A., Dell'Accio, A., Giambó, S., Naccarato, G., Pantano, P.: WebGIS and tourist personalized itineraries for exploitation of calabrian cultural and archaeological heritage. In: Proc. 2nd International Conference On Remote Sensing in Archaeology (2006)

15. Bertacchini, P.A., Dell'Accio, A., Mallamaci, L., Pantano, E.: Benefits of Innovative Technologies for territorial Communication: the Case of Study Virtual Museum Net of Magna Graecia. In: Proc. 5th Eurographics Italian Chater Conference, pp. 181–185 (2007)
16. Bertacchini, P.A., Reitano, A., Di Bianco, E., Pantano, E.: Knoledge media design and museum communication. In: Proc. 3rd International Conference of Museology (2006)
17. Vlahakis, V., Ioannidis, N., Karigiannis, J., Tsotros, M., Gounaris, M.: Virtual Reality and Information Technology for Archaeological site promotion. In: Proc. 5th International Conference on Business Information Systems (BIS 2002) (2002)
18. Mignonneau, L., Sommerei, C.: Designing emotional, metaphoric, natural and intuitive interfaces for interactive art, edutainment and mobile communications. Computer & Graphics 29, 837–851 (2005)
19. Mason, D.D.M., McCarthy, C.: The feeling of exclusion: Young peoples' perceptions of art galleries. Museum Management and Curatorship 21, 20–31 (2006)

Study of Game Scheme for Elementary Historical Education

Haiyan Wu and Xun Wang

College of Computer and Information Engineering,
Zhejiang GongShang University. Hangzhou 310018, China
why@mail.zjgsu.edu.cn

Abstract. It is important for development of education game to get the correct balance and find the best hybrid mode between education and game. In this paper, the factor considered during the design of education game is discussed, and several hybrid modes of education and game are proposed. Specifically for elementary historical education, the paper research on education game scheme of network role-play from the perspective of educational psychology, and design a case of historical education game using RPG mode with fully considering of knowledge and education. Players experience history by historical role-play, convey the sentiment of history by communication of players, and do with the hybrid integration of the game and education.

Keywords: education game, historical education, game scheme.

1 Introduction

With the expansion of the game industry, game had a great impact on the society especially for young people. Some primary and middle school students wallowed in online games and neglected their studies. Other purpose besides entertainment of computer games is increasingly concerned. The United States first proposed the concept of "Serious Game", computer games can be used in educational, training, simulation and other areas etc.

In many European countries, educational and puzzle games is developed earlier and gained a number of specific practical applications. In 1984, the famous American Electronic Arts (EA) was issued the "Seven Cities of Gold". One of the EA founder Trip Hawkins combined the two words "Education" and "Entertainment" together and create the new terminology "Edutainment"1. On the other hand, a number of websites for online education games abroad is very extensive. Relative to the maturity of education game abroad, there are few domestic enterprises engaged in education game. Shanda Interactive Entertainment Limited (SNDA) developed the first domestic youth-oriented education game which named "Learn from Lei Feng". This game has been widely concerned among researchers.

Most of the education game is developed for primary students, the majority subjects of education game are focused on mathematics, language courses, and many few on science, history, geography, physics, and other courses. This paper directed at

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 421–426, 2008.

elementary historical education, discussed the factors we should consider during the design of education game, and research the hybrid mode of game and education. Finally, a case of education game using hybrid mode is proposed. Students can learn historical knowledge by completing the task of the game.

2 About Factors of Education Game Scheme

Although most scholars think that computer games can be used in education to achieve the purpose of edutainment. But there are more difficulties in the design and application of the education game. The purpose of education limits the design direction of the game. The proportion of education and game has been the focus of controversy. Many educators believe that we should add large amount of educational content into games even at the expense of losing gameplay. Some people believe that education is a game first of all, the online game should funny players but not jaded, so not all of the knowledge are suited to be reflected in the games. We can only select the knowledge which is suited for designing to join into the game in a suitable manner. The design of the entertainment mode and innovation of education game is more difficult than ordinary online games. The following we propose a number of factors of education game scheme for elementary history education:

Respect History: Education theme of the game is rather difficult to grasp for history-oriented education game, the rules of the game is to allow failure, but some educational material is of historical significance, the results do not change. It seems a "bottleneck" that we should respect history but also have to consider the virtual of game. You can not change the fact that the emperor uniform china ultimately. And there is no way to make the Red Army crossed the grasslands unsuccessful. In light of this situation, Learners can choose its own position and decide their action in some major historical events in the education game.

Competitive & Collaborative: The process of playing game is an exploration and learning process curiosity-driven. We should try our best to retain game's competitive and challenging in the design of education game. So students can join into the learning process and acquire knowledge and skills when they play the game. Each player in the competitive game was going all out to upgrade, improve their equipments and gain a real sense of satisfaction and achievement. And we should also consider their collaboration in the game. Some tasks in the game can be designed to required teamwork. It provides communication environment in the process of collaborative. In addition, we can strengthen incentives mechanisms, the player would be awarded if they have completed the task excellent. This will encourage people to fulfill their tasks well and control the knowledge by the process of playing game.

Teaching-oriented: The purpose of teaching primary and middle school students is not just the pursuit of pure knowledge. Textbooks are no longer one, but more and more. Teaching is no longer just instilling mode, force-feeding, but to enable students to participate and explore, develop their analytical and problem-solving abilities, as well as exchanges and cooperation. Directed at the primary and secondary school education, we should give full consideration of teaching purpose in the design of the game. Education games should provide historical events for learners, and help people convey the sentiment of history.

Help System: Students may encounter some problems in the course of the game, such as the lack of understanding of the rules of the game, the knowledge embedded in the game and so on. They will not go on playing if they suspended because of the difficulties encountered in the game, and we can not achieve the purpose of education, So the game need to provide help system, Help information can provide from NPC, partners etc. Students can enhance and grasp knowledge of history through help system.

Information Feedback: Playing process is also a learning process, it is important to provide a feedback for improving the effectiveness and quality of learning. So, if we add the feedback system in education game, the student can understand their own shortcomings, it will help them to create confidence.

3 Hybrid Mode of Education and Game

It is the key for education game scheme to deal with the balance of education and game and find its hybrid mode.

3.1 RPG and Flash

It is the ultimate goal of the education game to integrate education naturally into the network games (especially MMORPG-online games). For historical education, it is relatively easily to map historical knowledge to game through some elements of the games such as tasks, skills, stags and so on. At present, education game can be divided into small Flash game and major RPG games, RPG game has complete story, unified modeling figures, but it is more difficult to develop, and it need a long life cycle and high costs. Flash game has no unified story and characters shape, it is more easy to develop, and it need low cycle and low cost. In general, flash game is suitable for young children, and primary and middle school students are more inclined large RPG game.

3.2 Online Knowledge Race

Knowledge can be merged naturally into the game is limited. So the second hybrid mode of education and game is knowledge race by online war. We provide a large-capacity database. The form of online war game can be responder, quiz, challenge champion and so on. The result of online war can also have an impact on currencies, decorations, equipment items, the value of experience, game time of player etc.

This mode is suitable for encyclopedia-knowledge which is interesting and widely adaptability, it can also link to the database of Entrance Examination and become an online learning platform of practice tests. Online war games can be set up in different districts for the local database and solve the problem of different versions of textbooks. We can provide the interface, so the local teachers can editor the local database according to local textbooks.

3.3 Educational Websites with Game

The third mode is combined with educational websites. We can put historical knowledge into the websites which can include the following modules: teacher talks, network courses, practice database, encyclopedia knowledge and so on. Educational websites

and online games have certain inherent link by using a unified website and game account. Learning in the educational websites will affect the game money, decorations, equipment items, the value of experience etc.

Website and the game may not have direct link, it can be only as a reference to the contents of the study, a reference to the answer for the game, a reference to complete tasks.

No matter which hybrid mode we choose, primary and secondary school students can learn academic knowledge, encyclopedia in the course of the game if we use proper treatment methods in the game scheme. We can also join the collaborating and explore learning mode into the game and affect ethical behavior of students on network. But the hybrid mode of game and education should keep to be researched before it is accepted by school students and their parents.

4 A Case of Historical Education Game Using RPG Mode

The following we design a case of historical education game using RPG mode. The Background of the game is designed as a prevalence magic Era, all the children would like to enter the magic school to study magic. Study in magic school is challenging but fun-filled. Magic school provides all possible ways of learning and athletic, in these ways, you can quickly grow. The architecture of game shows as figure 1.

Magic school is composed of the following models, "Trading System", "Global System", "Arena System", and "Magic Door".

(1)Trading System: Purchase of resources, exchange of resources

(2) Globe System: Students can understand the location knowledge by globe.

(3) Arena System: Gamers earn money by examinations or competitions.

(4) Magic Door: Magic Door is a RPG game. We chose 12 historical events and designed 12 scenes by them. These historical events include "Three Kingdoms", "The Silk Road", "Zheng He's Sailing to West Ocean" etc. Each scene contains a number of tasks, the screenshot of scene "Zheng He's Sailing to West Ocean" is showed as figure 2. Students can enter the history space and experienced the historical events from the Magic Door. If player complete the tasks successfully, he will have considerable incentives for an early realization of the spell dream. Various historical events were designed as a stag in RPG game. Students can learn historical knowledge by completing the task.

Magic school in all activities is integrated. It includes experience integration and property integration.

(1) Experience integration: Player is awarded by the level title when reach a certain level (0-10: Magician Trainee, 11-20: Magician Internship, 21-30: Junior Magician, 31-50: Intermediate Magician, 51-80: Advanced Magician, 81-100: Magician Master). Magician Master becomes "Magician King" above level 100.

(2) Property integration: Wealth can be achieved to a certain extent, "Magic Tycoon"

Magic guiding NPC is a "smart girl". She provides walkthrough of the game, such as the rules of game, notices, upgrading information etc.

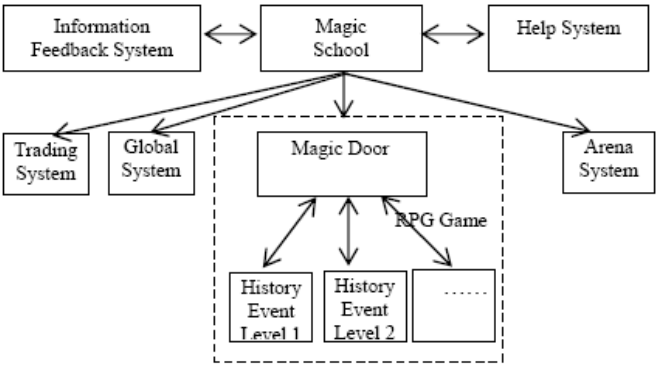


Fig. 1. The architecture of education game

Fig. 2. Screenshot of scene "Zheng He's Sailing to West Ocean"

5 Conclusions

The hybrid mode of education and game should be widely researched before it became an accepted educational means. This paper aimed at elementary historical education, and designed a case of education game using proposed hybrid mode. Students are absorbed with virtual role-play in the magic school and learn historical knowledge by completing the task. Meanwhile, the game also provides information feedback system, and helps a student understanding their lack of the study.

Acknowledgments

This work is supported in part by the grand foundation of science and technology department of Zhejiang province under grant No.2005C13021, Zhejiang natural science

foundation under grant No.Y105579, the key foundation of education department of Zhejiang province under grant No.20050642.

References

[1] http://game.86516.com/Game/news/chanye/130154474.htm
[2] Becker, K.: Teaching with games: The minesweeper and asteroid experience. Journal of Circuits, Systems and Computers 17(2), 23–33 (2001)
[3] Coyne, R.: Mindless repetition: Learning from computer games. Design Studies 24(3), 199–212 (2003)
[4] Durkin, K., Barber, B.: Not so doomed: Computer game play and positive adolescent development. Applied Developmental Psychology 23, 373–392 (2002)
[5] Squire, K.: The Games-to-Teach Research Team. Design Principles of Next-Generation Gaming for Education. Educational Technology 43(5), 17–23 (2003)
[6] Virvou, M., Katsionis, G., Manos, K.: Combining software games with education: Evaluation of its educational effectiveness. Educational Technology & Society 8(2), 54–65 (2005)

Integration of Game Elements with Role Play in Collaborative Learning — A Case Study of Quasi-GBL in Chinese Higher Education

Zhi Han[1] and Zhenhong Zhang[2]

[1] College of Software, Nankai University, Tianjin, P.R. China
[2] School of Educational Technology, Beijing Normal University, Beijing, P.R. China
hanzhi@nankai.edu.cn, zhenhong.zhang@gmail.com

Abstract. The reason that most undergraduate students in China spend a lot more time playing computer games rather than learning is not that learning is too hard, but it is boring, while playing games is fun. How to provide appropriate and interesting opportunities that can engage students and improve the learning process? This paper discusses an innovatory education paradigm quasi-GBL, which integrates game elements with role play in collaborative learning. Case study of quasi-GBL in an undergraduate course "Software Engineering" reveals that quasi-GBL is successful in achieving experiential and fun learning, offering a variety of knowledge presentations, and creating opportunities for knowledge application.

Keywords: quasi-GBL, role play, collaborative learning, case study, software engineering.

1 Introduction

In recent years, college teachers in China have been harassed by the too much engagement of undergraduate students in computer games. According to the 7th Online Game Research Report of China, 43% of the game players are from the age of 19 to 25, 56% of the players are undergraduate students, and the average time spent in playing games is 3 to 6 hours per day[1]. On the other hand, questionnaire in a university in China shows that 50.75% undergraduate students miss classes frequently and those undergraduate classes that over 80% of the enrolled students attend account for only 58.61% of the total[2]. The following questions are thus proposed: Why do undergraduate students engage so much in games rather than in learning? How to design effective learning opportunities to respond to this challenge? According to Prensky, what we are waiting for is the great teacher-designers to step forth, the people with the vision to harness games in the name of fun learning[3].

Analysis of game playing experiences indicates that games may offer supports in several aspects of the learning process: learners are encouraged to combine knowledge from different areas to choose a solution or to make a decision at a certain point, learners can test how the outcome of the game changes based on their decisions and actions, learners are encouraged to contact other team members and discuss and

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 427–435, 2008.

negotiate subsequent steps, thus improving, among other things, their social skills[4]. Games carry enormous potential to create immersive, experiential learning environments, draw students into a project and enhance their capabilities in information processing, decision making, knowledge application, problem solving, and group collaboration, which are still lacking with Chinese undergraduate students. Ignoring the educational power of games dismisses a potential valuable learning tool[5].

In this paper, the authors take initiative in facilitating and improving the learning process by integrating game elements with collaborative learning, and foster an innovatory education paradigm of quasi-GBL (quasi game based learning). It is 'quasi' because the pedagogy is not purely game based yet at the same time bears features of both games and collaborative learning. Quasi-GBL, offering immersive experience, is both engaging and effective for a broad spectrum of Chinese students. The authors apply quasi-GBL pedagogy to a hybrid undergraduate course 'Software Engineering' in Nankai University in China, which students report to be 'both fun and rewarding'.

This paper first discusses learning objectives and characteristics of the course 'Software Engineering'. Game Based Learning (GBL) is then examined and the obstacles to its widespread application into formal education are discussed. It is then revealed, through a case study, that quasi-GBL, an innovatory alternative to 100% GBL and more compatible with formal education, offers effective learning environments in the instruction of the undergraduate course 'Software Engineering'. The study concludes that time has come for games to be integrated with education and the key is to consider how best games can be used, for which quasi-GBL is a good example.

2 Instruction of Software Engineering in China

"Software Engineering" is a required course for both undergraduate and graduate students who major in subjects related to computer in most Chinese higher education (HE) institutions. The course is designed to present students with both technological skills and engineering rationales in the design, development, operation, and maintenance of software systems. Students are expected, after learning the course, to have sufficient knowledge in requirement analysis, software design and development, technical writing, and team work, so that they can enter employment at any position in the software lifecycle. The course, offering a broad range of knowledge and skills, goes beyond pure technology to encompassing technical, administrative, and social aspects, and calls for highly experiential (real or simulated) learning environments.

However, teacher centered pedagogy is still prevailing in instruction of the course in China. Students are supposed to learn the principles, steps, and theories by rote, without making sense or experiencing. Even those who have secured high scores in the course do not find the course useful and rewarding. As one student said when interviewed by the authors, "Software engineering is a very difficult course and it took me much time to memorize those boring stuff which I forgot as soon as the final exam was over."

Identifying such a problem, some teachers [6-8] try to offer better and more interesting learning options to students with case study method, in which teachers guide students in

analysis of a case of software development cycle and facilitate discussions. Case study integrated instruction engages students more by involving them in observation and exploration of real cases. However, according to Chris Dede, an ideal learning environment allows students to alternate between being "inside" an environment and being an outsider looking in[9], and case study only offers the latter. Active and effective learning in the course "Software Engineering" can only be facilitated when students are inside immersive environments into which knowledge and skills in the subject are built and which represent distributed professionalism in various roles. Quasi-GBL provides such learning environments.

3 Quasi-GBL (Quasi Game Based Learning)

Despite the enormous potential of GBL, it is still difficulty to integrate games into curriculum of formal education, because of the difficulty in identifying their relevance to the curriculum, potential educational benefits, and practical integration method[10]. Quasi-GBL, integration of game elements with role play in collaborative learning, is an innovatory education paradigm that exerts advantages of GBL and at the same time fits better into formal higher education in China.

3.1 GBL (Game Based Learning)

Games, especially computer games, are an important part of leisure lives for young people and are becoming a part of culture as well. The most popular computer games include role play games, real-time strategy games, shooting and fighting games, adventure games, action games, puzzle games, and chess games. Games are engaging for their features like interactivity [11], rules and goal[12], challenge[13], curiosity and control[14], etc., which hold educational potential if managed properly. In particular, games have a high learning value in certain educational domains that stress group communication and collaboration, decision making, and self exploration.

There has been a longstanding rift between games and more "worthy" activities, such as learning. It is only in recent years that people are interested in asking whether games can offer better learning environments and how to manage them. GBL (Game Based Learning) looks into ways to integrate games into learning, which may enhance students' intrinsic motivation, create immersive experiences, involve social interactions and collaboration. However, with positive results of quite a few studies on application of games to teaching and learning, university teachers still find it difficult to integrate games into their instruction, mainly because game based learning does not fit in with formal university education and there are few educational games available for university students. Thus an innovatory alternative to GBL—quasi-GBL, is proposed by the authors, which may exert the educational potentials of GBL and also be applied to formal undergraduate education in China.

3.2 Quasi-GBL and Instruction of "Software Engineering"

Quasi-GBL (Quasi Game Based Learning), a term coined by the authors, refers to the instructional method that integrates game elements with role play in collaborative

learning. It can be applied to face-to-face (F2F) classes, where quasi-GBL takes place in a classroom or to hybrid classes, where it relies partly on a virtual learning environment (VLE). In this study, quasi-GBL is used in the undergraduate hybrid course "Software Engineering" at Nankai University in China. The course is hybrid in the sense that some teaching and learning is done in the classroom while students can also communicate with each other, share learning materials, and engage in collaborative production on an online learning platform.

Role play finds its application in the course as it allows students to 'be' certain roles that would otherwise be inaccessible to them, thus experiencing the ways a certain type of role thinks about and solves problems. Another characteristic of role play is that students have to work collaboratively with others and practice social and communication skills. Role play holds special educational value in the course "Software Engineering" because it distributes expertise in the software development cycle among the roles, requiring students to learn the skills and collaborate with others in a team. Yet the prevailing role play model in China, often used in language classes and highlighting the stage performance of players in different roles, does not fit in well with tasks in the course "Software Engineering", which takes much more time and involves scheme proposing, trials, arguments, decision making, problem solving in group collaboration, while display of the product is only a part of the prolonged process. In order to engage students in the complicated and prolonged tasks and retain their interest and motivation in the process, game elements are integrated with role play in collaborative learning.

Seven basic elements are identified in games, including goal, rule, competition, challenge, fantasy, safety, and entertainment[15]. In quasi-GBL, these elements penetrate into role play and manifest themselves in forms of real problems, individual tasks and group collaboration, scores, puzzles, awards, and replays. In the course "Software Engineering", students work in small groups simulating software development teams or companies, each playing a certain role. The teacher acts as a client to each team and starts the role play by proposing several "Easy level" tasks for groups to choose from. Groups will also be given puzzles with various awards at each of the four stages in the software development cycle: Requirement analysis, Design, Implementation, and Deployment. Successful solving of a puzzle can win extra scores for the group. At the end of the task each group presents the product and related documents to the class, which will be assessed by the class and the teacher with an eye to the original requirement. Besides, each student submits a report to the teacher evaluating his group members' and his own performance in the task, which will also be counted in the assessment. Then another round of role play starts with the students changing roles or groups and the teacher, that is, the client, proposing "Normal level" tasks. The process of quasi-GBL application to the course "Software Engineering" is shown in Figure 1.

Quasi-GBL adopts a holistic approach to learners' assessment: class assessment of the group product, teacher's assessment of the group product, and students' evaluation of group partners and themselves all count in it. A case study of quasi-GBL in the undergraduate course "Software Engineering" is presented in the following part.

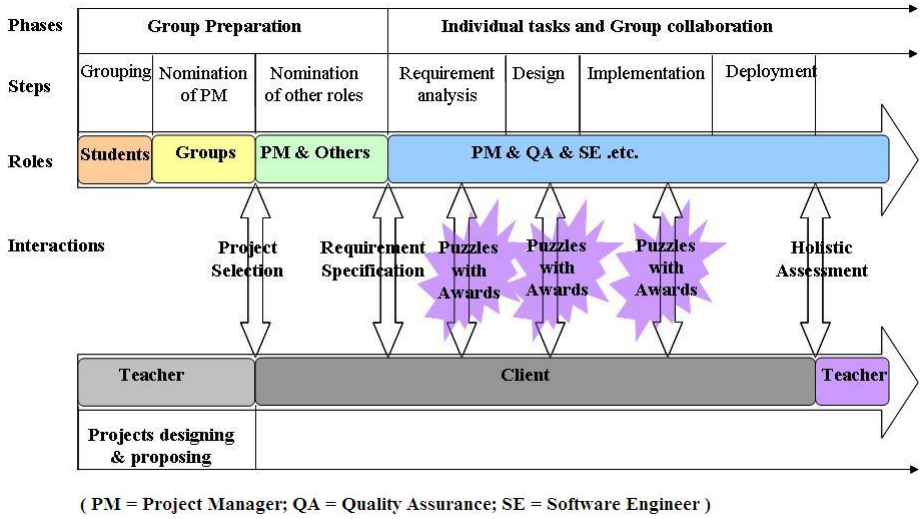


Fig. 1. Process of Quasi-GBL application in "Software Engineering"

4 Case Study

4.1 Course Context

The undergraduate required course "Software Engineering", lasting for one semester at Nankai University, is a hybrid one in which there is a 3-hour face-to-face session every week and a virtual learning environment supporting online communications and discussions about tasks and problems, and sharing of learning materials. 150 undergraduate computer majored students in their third year, from the age of 19 to 22, enrolled in the course and are divided into three classes. Quasi-GBL is applied in one of the three classes for a semester by one of the authors as the course coordinator and lecturer.

4.2 Quasi-GBL Instructional Design and Implementation

At the beginning of the semester students were asked to form nine groups with five members in each and a project manager was selected in each group. The teacher, as client, started the role play by putting forth five 6-week "Easy level" software development projects for groups to choose from. Take Group 5 for example. The project chosen by Group 5 said, "Please develop an online purchasing system for our company in 6 weeks. For further details as to its functions and other requirements, please contact the client". Project Manager 5 then organized a group meeting and guided in role assignment within the group: three were software engineers and the other was quality assurance. According to what they had learned in face-to-face sessions, Group 5 contacted the client (the teacher) and conducted requirement analysis in the first week, through one face-to-face meeting and three emails. At the end of the second week, the client put forward three puzzles for the group: Puzzle 1 was an additional

speed requirement for the system with an award of 30 points to the total score of 100 points for the product; Puzzle 2 was a compatibility requirement for the system with an award of 10 points; Puzzle 3 was an additional interface requirement with an award of 20 points. Project Manager 5 organized another group meeting and there was a heated discussion on how to cope with the three puzzles in the existing development scheme. The project manager insisted on trying to solve all the puzzles while the three software engineers thought it impossible. Finally the project manager was isolated and only Puzzle 1 was included in the modified development scheme, which was later successfully solved. Group 5 went into the stage of system design from the beginning of the third week, which took much longer time than expected, and it was in the middle of the fifth week when they started coding. It turned out that at the end of the sixth week when the system was distributed there was a function of "Rating the purchaser by its purchasing record" missing. In the process other puzzles were given and another puzzle with 20 points' award was solved. In the seventh week for software presentation, Group 5 obtained an average score of 122 (85 for the product and 37 plus for the puzzles) from the class and 124 (83 for the product and 41 for the puzzles) from the teacher. Besides, the group members' performance evaluation reports were submitted, from which it was found that the project manager failed in organizing the team work and was almost isolated in the whole process, while one of the software engineers not only contributed much more than the others and solved the two puzzles all by himself but also did a lot of management work for the group, which was taken into consideration in the assessment of each member in the group. Then another round of role play started with the client proposing five 9-week "Normal level" projects at the beginning of the eighth week and each group changed roles and chose a project. The "Normal level" projects were more difficult than "Easy level" ones yet the role play went on in a similar style, which was an opportunity for students to "replay" in another role.

At the end of the semester, focus groups interviews were conducted to evaluate the effectiveness of quasi-GBL pedagogy. Patton suggests that focus groups interviews are essential in the evaluation process[16] and selection of interviewees is essential for the rigor of the evaluation. According to Stewart and Shamdasani, the group must consist of representative members of the larger population[17]. In this case, two group interviews were conducted: one with 6 students in the class and the other with 9 students. The authors first sub-grouped the class with reference to their participation and activeness in the course and then carried out convenience sampling in each of the subgroups. 6 from the more active participation sub-group and 9 from the less active one were selected as interviewees. Each focus group interview lasted about an hour and 5 to 7 open-ended questions were asked about the students' experiences in the course with quasi-GBL. The first author acted as interviewer and the interviews were recorded, transcribed, and then analyzed by the two authors. Analysis of the interview transcripts reveals that students are mainly impressed by two features of quasi-GBL: experience and fun.

According to the interviews, students believe that they have learned more from experiences than from listening to the teacher in classrooms or discussing with their classmates in case studies alone. First, they have experiences in collaboration. In quasi-GBL, students need to assign tasks among the group, communicate with partners face-to-face or online, practice team administration and team work, and try to

solve conflicts and problems, from which they can learn a lot about techniques and skills in collaboration. Second, they have experiences in the real process of a software development cycle. Immersive experience in the course projects has helped students to better understand what the tasks and focus are in each stage of the process, how the development work is organized and implemented, and how complicated the process might be. Third, they have a better understanding of the diverse roles involved in software lifecycle, including required competencies and duties for each role. More important, some students have identified their potentials, weaknesses, and preference through quasi-GBL, which is a crucial part of holistic development.

Aside from experience, fun is a peculiar feature of quasi-GBL, which engages students and makes them more active in learning. According to the interviews, fun elements come mainly from collaboration, role play, and puzzle solving. Most students enjoy social interactions with group partners and draw much fun from sharing of ideas, information, problems, and success. Role play also offers fun because it helps students out of the boring memorization and they can complete learning through practice and experience, in which there are a variety of real decisions to make, problems to solve, and responsibilities to take. Puzzle solving, with challenges and the possible award, is reported to be the most fun in quasi-GBL. As one student said in the interview, "Puzzles are really engaging and I can spend hours thinking about it and trying to solve it, because it is sort of a test on how capable I am."

4.3 Discussions

With puzzles added to the role play, students report more fun and motivation in learning, as one student said, "I am more active in learning than before and every member in our group feels the same. We like puzzles, and it's challenging and fascinating to solve them by ourselves." In solving the puzzles, students access resources, apply related knowledge they have learned, and engage in collaboration and communications. They willingly spend a lot of time on puzzles, just like when they are playing a game.

Score element in the role play is another incentive to students. After the seventh week when their products were scored by the class and the teacher, students had a higher sense of achievement and social belonging, as one student wrote in his performance report, "We were excited on the day when presenting our group product to the class and more excited to know how many points we'd got for it. It's different from knowing my score of the exam in other courses, as this was the joy that I could share with my partners at what we'd been working on for the last several weeks. And the score also gave us motivation to work harder and achieve higher marks in the next project". Some have complained about Chinese students' caring about their marks in exams more than anything else, yet score element in the role play has re-directed their attention from memorizing information for the exam to making a better product and solving the puzzles, in which process real learning happens while they are analyzing, arguing, and trying. Scoring other group's products is also a significant experience for students. They can learn a lot from looking at other groups' products, as one student said in the interview, "Learning how others solve problems that I have little idea of is a wonderful experience, which gives me a lot of insights and inspiration". Besides,

the responsibility of assessing other groups' products makes students more active in learning as they feel they have their say in learning.

A problem students have in team work is how to deal with different opinions and ideas without impeding the completion of tasks. In Group 5 there was so heated argument between the project manager and the system engineers that later on the project manager was isolated and the system engineers communicated with the client direct and completed the task by themselves. In the performance report, the system engineers complained about the project manager's "not listening to others, arbitrary, and not cooperative". The interesting part of the story is that in the "Normal level" project when the former project manager became a system engineer, he seemed to have learned something about team work as he wrote in his performance report about his own alteration, "I've understood how my partners felt in the first project because now I'm a developer. They need my suggestions, not orders". Team spirit and capability in collaborating with others are part of essential qualities for people in general, and more in software development industry as the division of labor in the software development cycle is becoming ever delicate. Experience in the role play can help students to identify their strengths and weaknesses in team work, which is the starting point to improvement and self development.

A special feature of the collaboration in the case is that it is hybrid, that is, both face-to-face and online. Online collaboration is different from F2F collaboration, for lack of real time feedback, gestures, expressions, and body language, just to name a few. It is interesting, thus, to notice that in the course, most communications online were via chatting systems while emails or online forums are seldom used between students. When asked for the reasons, students said that they'd tried to use online forums yet found them to be so "inefficient". With software development becoming more and more globalized, the ability to work collaboratively online with partners is a necessity for future professionals. Immersion in such an experience is an important benefit for the students.

5 Conclusions

Game based learning has not found many applications in schools or universities, universities in particular, because it does not fit in well with curriculum of formal education and schools are skeptic about whether it brings more benefits than harms. Quasi-GBL, an alternative to GBL, which integrates game elements with role play in collaborative learning, is an innovatory education paradigm in realizing "fun learning" and brings broad learning benefits to students. In the undergraduate course 'Software Engineering' studied in this paper, it is found that students enjoy the experiential learning process, and more important, they are engaged in learning more, because learning becomes more fun. Students have also benefited in a broader spectrum through quasi-GBL, in aspects of collaboration and team work both face-to-face and online, communications with people, utilization of information for problem solving, planning and decision making, not by memorization but through experience. Quasi-GBL can also be applied to other courses in higher education, whether a face-to-face course or a hybrid one, and more studies need to be carried out in such field.

References

1. iResearch Consulting Group.: The 7th Online Game Research Report in China. iResearch Consulting Group, 9-12 (2007)
2. Wen, K.: Countermeasure research and cause analysis of the existing college students' study-weary phenomenon. Journal of Tianjin University of Technology and Education 12, 58–61 (2006)
3. Prensky, M.: Digital Game-Based Learning. McGraw-Hill, New York (2001)
4. Pivec, M., Dziabenko, O., Schinnerl, I.: Aspects of game-based learning. In: I-KNOW 03, the Third International Conference on Knowledge Management, Austria, pp. 178–187 (2003)
5. Oblinger, D.G.: Games and Learning. Educause Quarterly 3, 5–7 (2006)
6. Wang, X., Wang, J.: Case Study in Software Engineering Instruction. Computer & Information Technology 6, 114–118 (2006)
7. Wu, H.: Application of Case Study in Software Engineering Instruction. Journal of Yichun University 4, 98–99 (2007)
8. Ye, J., et al.: Research into Case study in Software Engineering Instruction. IT Education 7, 19–21 (2007)
9. Dede, C.: Planning for Neo-Millennial Learning Styles: Implications for Investment in Technology and Faculty. In: G.D., Oblinger, J.L. (eds.) Educating the Net Generation, EDUCAUSE, Boulder, Colo (2005)
10. Kirriemuir, J., McFarlane, A.: Literature Review in Games and Learning. A Report of NESTA Futurelab, `http://www.nestafuturelab.org/research/reviews/08_01.htm`
11. Tornton, G.C., Cleveland, J.N.: Developing managerial talent through simulation. American Psychologist 45, 190–199 (1990)
12. Johnston, R.T., Felix, W.: Learning from video games. Computer in the Schools 9, 199–233 (1993)
13. Baranauskas, M., Neto, N., Borges, M.: Learning at work through a multi-user synchronous simulation game. In: Proceeding of the PEG 1999 Conference, pp. 137–144. University of Exeter, Exeter, UK (1999)
14. Malone, T.W.: What makes computer games fun? Byte 6(12), 258–277 (1981)
15. Alessi, S.M., Trollip, S.R.: Multimedia for Learning: Methods and Development. Allyn and Bacon, NY (2001)
16. Patton, M.Q.: Qualitative evaluation and research methods, 2nd edn. Sage, London (1990)
17. Stewart, D.W., Shamdasani, P.N.: Focus groups: Theory and practice. Sage, London (1990)

A Case of 3D Educational Game Design and Implementation

Huimin Shi, Yi Li, and Haining You

Center for Research on EduGame, Nanjing Normal University, Nanjing, 210097, China
nj_huizi@sohu.com

Abstract. Games are increasingly used for educational purpose. Player can obtain experiences through playing game. Game is a safe and efficient way for security education. This paper introduces a game design: *Escapre from Fire*. a design procedure and system structure are outlined. Related programming technologies are discussed.

Keywords: Educational game, Security education, Design procedure.

1 Game and Learning

Every year, millions of people die of disasters, including traffic accident, electric accident, fire, flood, earthquake and so on. Disasters bring people endless sadness and loss. However, most of the dead should have survived if they had known more about security technique. It is not possible to train people in real situation. Up to now ,there's no sucessful instructional model of security eduction in china. It is necessary for us to find out some safe, realistic and efficient solution. Games can provide us a good and safe way for security education. The educational potential of computer games is often celebrated.The structure of activities embedded in computer games (as opposed to the game content) develops a number of cognitive skills[1]. Engaged in repeated judgement-behaviour--feedback loops[2], game players can obtain experence which is helpful for dealing with real disaster.

2 Related Work

2.1 911: First Responsible

Some fire games have been developed. In *911: First Responders*（Fig. 1）, the player acts as a fire commander in a fictional rescue and catastrophe management organization. The player can command a number of vehicles and staff to deal with all the accidents. For example, while the firemen are fighting with the fire, the doctors are required to cure the wounded. The command and the firemen can accumulate their reputations to become the top fire department [3].

2.2 Fire Captain

Fire Captain: Bay Area Infer（Fig. 2） is a fast-rhythm game. Based on two true fire accidents. One is about fire of Holand firework factory in 1993, and the other is

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 436–441, 2008.

Fig. 1. Scene of *911*

Fig. 2. Scene of *Fire Captain*

Auckland forest fire in 1996. The role of the player acts as a fireman to deal with all of the emergences [4].

3 Design of *Escape from Fire*

3.1 Game Analysis

Fire happens continually. There are 6~7 millions fires every year in the earth, more than 70% of which are house fire. In china, nearly 20000 millions RMB per year is lost because of fire and more than 2,000 people die of fires. We are developing a 3D fire game *Escape from fire* to teach children and the teenage how to escape from fire.

Different from *911* and *Fire Captain*, the player act as victim in the fire spot, not the fireman. From educational perspective, we summarized some features of this game, as the following shows.

Realistic style. It means the results form the player's operations accord with facts. The game is based on real fire cases, such as Tokyo Star 56 Mansion Fire in 2001, the Sinkiang Kelamayi Fire in 1994 , China and others from Discovery Program.

Time limit is another key character of this game. The fire will spread quickly and the player is given limmited time to finish the task.

Feedback. The players may fail during playing. Once they fail, the information about the wrong operation will come out.

Form game perspective, challenge, fantasty, curiosity and are the main features. [4] Unlike some eductional game, in which the knowdge is designed as questions and the only chanlenge for the play is to answer questions. This game is well structed , the player has to make decidions and take actions to sovle problems. Thus, the player construct their knowledge by exploring[6]. That is learning by doing, not by answering questions.

3.2 Design Procedure

How to combine education in game without lose of fun? It is a main difficulty for us, maybe for all of those educational game designers. Although there is no unique guideline for educational game design, all of them begin with instructional analysis, shown as fig. 3.

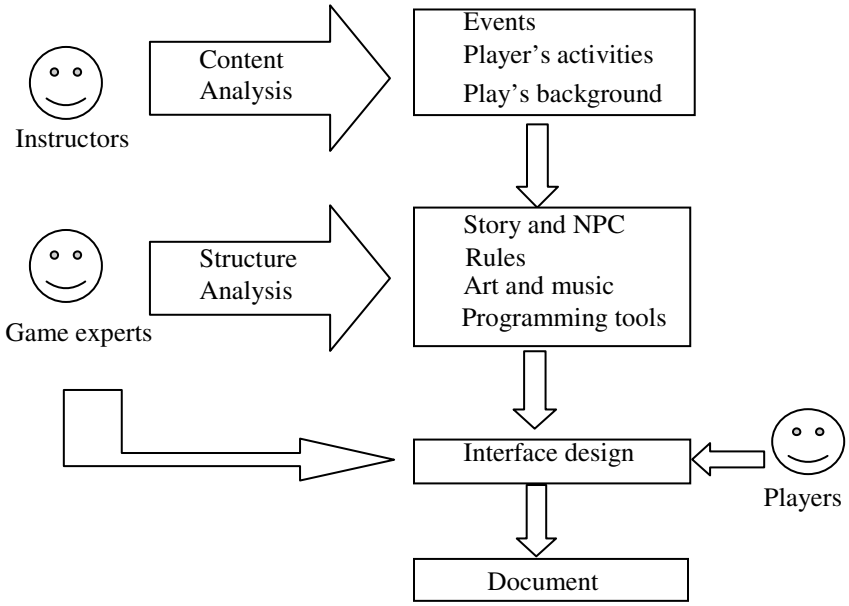


Fig. 3. Game design procedure

3.2.1 Content Analysis

Game design can be devided into two stages: content design and structure design. Educational experts focus on content analysis. During this game design, 2 experienced firemen were invited to join the team. The educational experts break fire knowledge into pieces as cut-in points for game design, as table 1 shows. Game style is also decided in this step.

3.2.2 Structure Analysis

Fantasy can be very important in creating intrinsically motivating environments. However, these must be carefully chosen to appeal to the target audience[7]. Instructor focus on what content the player shoul master after playing, while game disigner

Table 1. Content Analysis

Scene	Fire happens	Player's activities
House	Incorrect operation on electric iron Oil pan fire Old or exposed electrical wire Cloth sofa burned by electric heater	Put out the fire with correctly if possible Escape from fire and ask for help in time Use bedsheet, curtain and rope to escape from the veranda
Skyscraper	Self-igniting of paper sundries ripple effect of electric welding in the lift	Use fire extinguisher Emergency exit ,not the lift No panic, escape orderly Separate the fire out the door and window
Cinema	Set fire purposely by terrorists Gauze curtain burned by strong light	No panic, escape orderly Cover mouth with wet washrag

cares about what makes a computer application fun to operate. Based on the result of content analysis, game's story, characters, style, rules and programming tools etc. will be discussed in this step. This game is a strategy-based game, in which the player acts as a boy Dingding. He and his friend Xiaoqiang who is a clever but timid boy will encounter different fire. They have to put out the fire or escape from the fire. The OGRE game engine is decided.

4 Engine Analysis

ORGE is an open source game engine. It is mature, steady, credible, flexible and it can produce most of the effects that other popular engines can do, such as object, HDR, knaggy texturing etc,. Different from other engines, scene graph is separate from content objects (Fig. 4). Thus , it is simple to produce engine-based realtime 3D knaggy animation.

All of this geometry and these rendering properties are made available to the scene graph via MovableObject. Instead of subclassing the movable object from a scene node, it is attached to the scene node[8]. It is not necessary to set up the relationships among the scene nodes, and know about the details of the renderble objects. Expected functions can be carried out through operating on the scene graph interface. The scene graph interface can even change completely and the content classes would not be affected in the least.

How to produce real time animation in OGRE. First, we should built 3D models and animation in 3D modeling software(Maya, 3Dmax, etc,.). Second, these 3D models and animations are exported as *.Mesh and *.skeleton files with proprietary exporing tool of OGRE. Then, these *.Mesh and *.skeleton files are attacted to the scene nodes. Thus, the realtime animation can be created in OGRE.

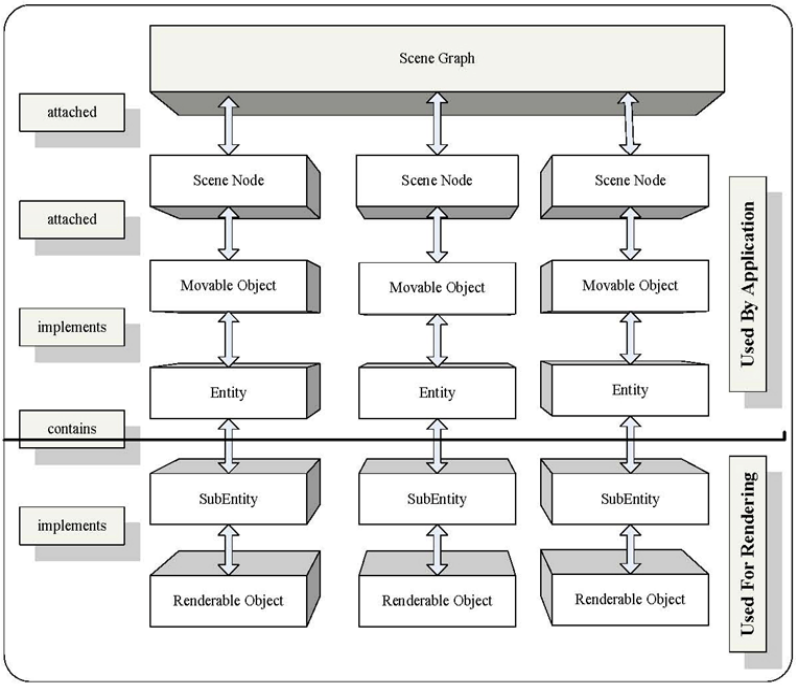


Fig. 4. Relationship between scene graph and content objects in OGRE

The following code is about how to add a node in mSceneMgr and attach a existing model called jaique to the node. The final rendering is shown as Fig. 5

```
Entity* obj01 = mSceneMgr->createEntity("R1",
 "jaiqua.mesh");
SceneNode* D1 = mSceneMgr->getRootSceneNode()-
 >createChildSceneNode();
Data01->attachObject(obj01);
```

Then, we set the playing model as spline interpolation, shadow effect as “true”, object’s emergencing coordiate, object’s circumvolving angle and axis in CreateSkeletonAnimation(). Then, animation information in jaiqua.skeleton is added to the nodels. Sneak actions form jaiqua.skeleton is given to jaiqua model.

```
void CreateSkeletonAnimation()
{
Animation::setDefaultInterpolationMode(Animation::IM_SP
 LINE);
Entity *ent = mSceneMgr->createEntity("M1",
 "jaiqua.mesh");
ent->setCastShadows( true );
 SceneNode* BJmovie = mSceneMgr->getRootSceneNode()-
 >createChildSceneNode("R1",Vector3(3,0,-3));
BJmovie->attachObject(ent);
```

```
BJmovie->yaw(Degree(-90));
mAnimState = ent->getAnimationState("Sneak");
mAnimState->setEnabled(true);
}
```

Fig. 5. Realistic animation in OGRE

5 Future Work

The games are still on the tentative stage and need the further improvements in technology(modeling techniques, interaction technology etc.) and designin. We are to improve the system so that it can provide three alterative game levels----from level 1 to level 3. The player can select game level by their own. For those experiened players, level 3 may be suitable, being most complex and challenging. For children and some girl players, level 1 may be suitable, bejing simple and less challenging.. In addition, we will try these games in schools and do some research so that we can evaluate and modify the games from player perspectives. It is urgent for us to attach importance to security education. We are planning to develop a series of games on security education for children, such as traffic and electricity

References

1. Robertsona, J., Howells, C.: Computer game design: Opportunities for successful learning. Computers & Education 50(2), 559–578 (2008)
2. Garris, R., Ahlers, R., Driskell, J.: Games, motivation and learning: A research and practice model. Simulation and Gaming 33(4), 441–467 (2002)
3. http://www.atari.com/us/games/911_first_responders/pc
4. http://www.firecaptain-game.com
5. Malone, T.W.: What makes things fun to learn? Heuristics for designing instructional computer games. In: Proceedings of: 3rd ACM SIGSMALL symposium and the first SIGPC symposium on small systems, pp. 162–169 (1980)
6. Liang, T.: A 3D Escape-from-Fire System. Journal of Shenyang Agricultural University (Social Sciences Edition) 9(5), 779–782 (2007)
7. Ebner, M., Holzinger, A.: Successful implementation of user-centered game based learning in higher education: An example from civil engineering. Computers & Education 49(3), 873–890 (2007)
8. Junker, G.: ProOGRE 3D Programming, p.39 (2006) ISBN-13: 978-1-59059-710-1

Mathematical Education Game Based on Augmented Reality

Hye Sun Lee and Jong Weon Lee

Mixed Reality and Interaction Lab, Sejong University,
98 Kunja-dong, Kwangjin-ku, Seoul 143-747, Korea
cariome@hotmail.com, jwlee@sejong.ac,kr

Abstract. The computer game industry has been grown rapidly due to the development of graphics and communication industry. The computer games have been taking on various forms with its development. Currently, the computer games are not only used for amusement but also used for other purposes. The computer games that try to provide users more than fun are called serious games. Many researchers have developed serious games for various fields such as education, medical treatment. Education especially stands out among those fields. This paper suggests a mathematical education game developed using Augmented Reality. It is a board game for kindergarten and elementary students. Augmented Reality is used to extend user's experience and to increase the usability of the system.

Keywords: Augmented Reality, Serious game, Edutainment.

1 Introduction

Computer games have been gained a lot of attention nowadays. Computer games have become one part of the culture. Currently, computer games are not only developed for simple amusement. Some developers try to obtain something more than just simple fun. These games are developed to train users the specific skills while enjoying them. These types of games are called serious games [2]. Serious games are also called functional games or social impact games, and their focus is to provide users with extra benefit and fun.

Especially educational games are gained huge attention from developers and users. There are many supports from schools, corporations, and institutes, and many researchers develop educational games in the various fields [1].

Educational games that combined learning and playing are very practical, and they become new entertainment contents. Already there are various ways to present contents that deal with educational games. The reason these educational games are growing is not because they are games that combine enjoyment and education, but because they are recognized as having beneficial contents. Edutainment is already being used in variety of fields for educational purposes, such as educational games, livelihood training, and politics education.

The purpose of this paper is to introduce an educational game using Augmented Reality (AR) technologies. AR is used to extend user's experience and the usability of

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 442–450, 2008.

the game. The board creation tool is also provided to increase users' interest by helping them to create their own contents.

The organization of this paper is composed as stated. Section 2 introduces related research areas. Section 3 explains the technical elements and practical usage plan of the proposed educational game. Section 4 concludes with a description of future plans.

2 Related Research

Research regarding educational game has been done actively, so many new educational games are developed. Here we only deal with online educational games and AR-based educational games, which are mostly related to the proposed game.

2.1 Online Educational Games

ALEPH Project Game is a good example of an educational game (Figure 1) [3]. Here, the matrix of English words moves its position, and players must guess the correct word using 12 different attacking options. This game requires strategic techniques from players, and players can beat the opponents by using turn attack. Through this game's process, players can train educational effect by practice English words repeatedly.

(a) (b)

Fig. 1. ALEPH Project (a) Fight scene (b) It is a scene that make word as alphabet

Another game is Power Politics. This game is a simulation game regarding American presidential election (Figure 2) [11]. Through this game, players can experience a simulated president election and learn about the democratic government system. Players select the presidential candidate that actually existed, and manage influential movement from the beginning of the election campaign to the end. Also, players generalize the election campaign and must come up with a strategy plan to get the candidate elected. Players can learn the view of democratic government and the election process through this game.

In addition to these games, there is "Food Force"[12], which allows players to experience emergency food formation via the World Food Program. There are many commercial educational games such as "Brain training" [13] by Nintendo DS. As referred to research carried out by the Britain Ministry of Education, simulation or

(a) (b)

Fig. 2. Power politics (a) Candidate selection (b) Manage election campaign

adventure games (Simcity, Rollercoaster Tycoon, etc.) develops children's strategic thoughts and planning capability. The Britain Ministry of Education also acknowledged the educational value of these games [14].

2.2 AR-Based Education Games

AR is a good environment for serious games. AR can maximize the player's experience. As a representative example, there is "Solar-System and Orbit Learning in Augmented Reality System" (Figure 3) [4]. This system promotes knowledge of volcano eruptions, relations of the solar system, information of Earth's surface and others through a camera.

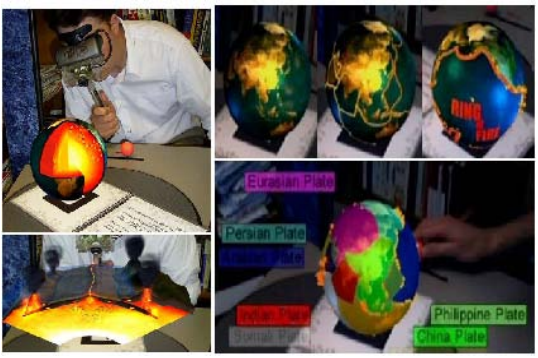
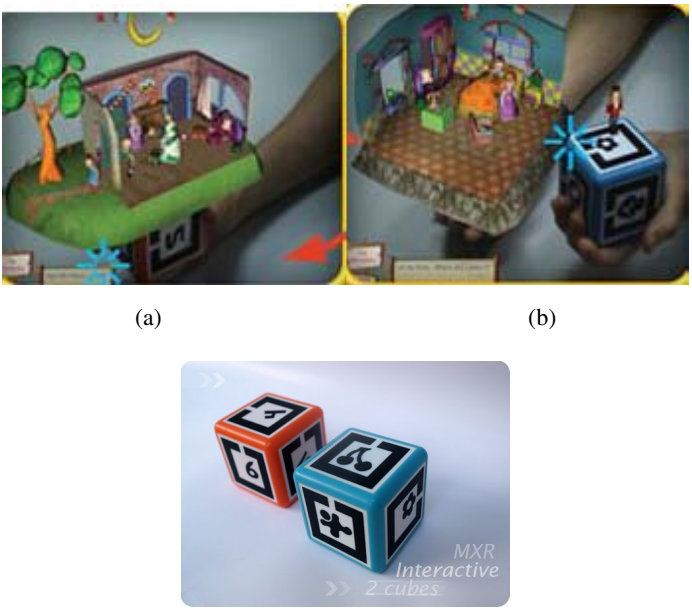


Fig. 3. Solar-System and Orbit Learning in Augmented Reality System

wlzQubesTM is a story book, which uses storytelling techniques for children (Figure 4) [5] [15]. Here, two cubes are used as a set. The first cube contains the main pictures of the story on 6 sides, and the second cube contains variety of needed items. Children roll the cube and correctly match the item to the scene of the story, which are viewed through the camera; and then the player tries to put the two cubes close to each other. The virtual items that are brought close together join together with the scene of the story, and the corresponding 3D animation pops up.



(a) (b)

(c)

Fig. 4. wlzQubesTM (a) Main scene (b) Selected item (c) Two cubes

The Cooking System [8] teaches cooking recipes, and AR Squash Game [6] combines AR and a sports game. Geometry System [7] allows players to use the virtual interface to control the size of the shape, move, and combine, separate, and so on.. Chemistry education System [9] combines and separates atoms chemically. These systems are combined with education in a broad and variety of ways.

3 Proposed Educational Game

The proposed educational game is the board game which can be used by kindergarten and elementary students. Augmented Reality is used to increase the usability of the game. The game is developed based on ARToolkit [10].

3.1 Scenario of the Game

The scenario of the proposed game is created based on the book Ria's Math Play [16] (Figure 5). This book introduces the concept of addition and subtraction to young children through playing a dice game of two players. The dice game uses three dice, (two numbered dice and one operator dice) and the board. Players roll the dice and calculate the outcome, and move a piece on the board according to the outcome. The player who arrives at the finishing point first wins the game. Since the dice game requires computation, parents or another helper has to play together with the children to teach the mathematical concepts. Children can only play with one board, potentially

leading to boredom in children. To overcome these limitations, we developed a computer game based on the book (Figure 6), but the computer game could not provide intuitive interactivity to children. Some parents also considered the computer game was unsuitable for young children. Because of these limitations of the board game and the computer game, we developed the AR-based board game based on the book, so children can enjoy the board game with various contents with the same intuitive interaction as the board game.

Fig. 5. Ria's Math Play

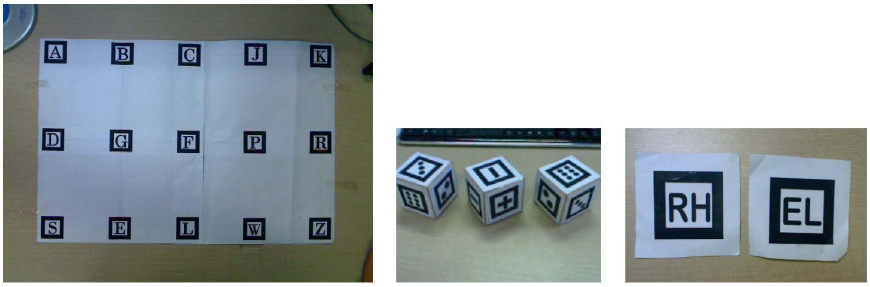
(a) (b)

Fig. 6. Computer game based on Ria's Math Play (a) Initial scene (b) Show scene that derive from value of three dice on the board

3.2 Proposed Mathematical Education Game

The proposed mathematical educational game targets young children. We assume the proposed game will be played in the living room with the TV, the computer and the PC Camera. The proposed game is played with three dice and the board made of markers (Figure 7). The board is created using multiple markers, so the system can augment virtual images even though a few markers are occluded by pieces of the game or players.

The number of markers used for the board is 15, which is decided by the viewing angle of the camera and the distance between the camera and the board. The size and the number of the markers are decided for reliable tracking based on the setup in which the game is played on.



(a) (b) (c)

Fig. 7. (a) The board with multiple markers (b) The operator dice and the numbered dice (c) Player's pieces

A different board can be selected according to scenario and user's preference among the given boards, and player's piece are augmented as 3D models on the selected board (Figure 8).

Fig. 8. The boards and player' s pieces

To use operator and numbered dice, the system has to distinguish the top marker among multiple markers of dice viewed by the camera. The system distinguishes it based on the 3D coordinate system between camera and marker shown in Figure 9.

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} & T_1 \\ R_{21} & R_{22} & R_{23} & T_2 \\ R_{31} & R_{32} & R_{33} & T_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_M \\ Y_M \\ Z_M \\ 1 \end{bmatrix} = T_{CM} \begin{bmatrix} X_M \\ Y_M \\ Z_M \\ 1 \end{bmatrix} \tag{1}$$

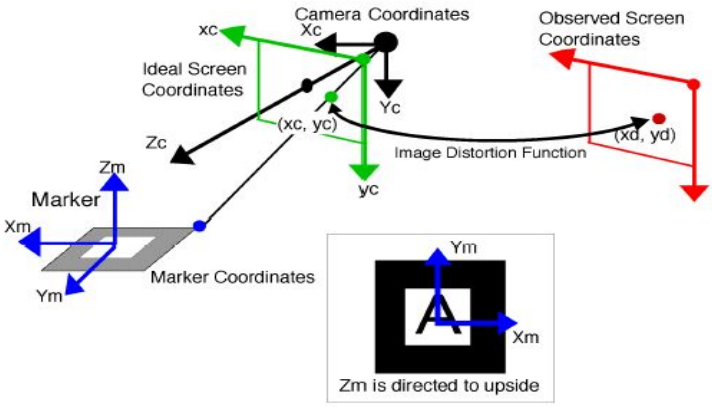


Fig. 9. 3D coordinate system between a camera and a marker

In Figure 9, X_C, Y_C, Z_C represents the camera coordinate, and X_M, Y_M, Z_M represents the marker coordinates. Through Equation (1), translation and rotation between the marker and the camera can be computed.

We compute the translation matrix T_A between each viewable marker on the dice and the camera and the translation matrix T_B between the board and the camera. The rotation and translation motion between the board and each viewable marker on the dice is computed using Equation (2).

$$T_A = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad T_B = \begin{bmatrix} B_{11} & B_{12} & B_{13} & B_{14} \\ B_{21} & B_{22} & B_{23} & B_{24} \\ B_{31} & B_{32} & B_{33} & B_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_{AB} = T_A^{-1} \cdot T_B = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} & B_{13} & B_{14} \\ B_{21} & B_{22} & B_{23} & B_{24} \\ B_{31} & B_{32} & B_{33} & B_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad T_{AB} = \begin{bmatrix} R_{11} & R_{12} & R_{13} & T_x \\ R_{21} & R_{22} & R_{23} & T_y \\ R_{31} & R_{32} & R_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} R_{3X3} & T_{3X1} \\ 0 & 1 \end{bmatrix} \tag{2}$$

The rotation matrix is presented as the XYZ method [17] in the Euler Angle expression. Angle distances between Z axis of each viewable marker of the dice and the board are computed, and one with the smallest angle distance within the given threshold is chosen as the Z axis of the top marker of the dice (Figure 10). Using this information, the system computes the outcome and overlay the corresponding information on the dice.

The system also provides a tool allowing users to create their own board and player's pieces to improve game's usability. The board is created by combining elements of the tool. Users can place a square, circle, polygon, straight line, rectangle, ellipse or other provided elements and can fill each component up with the given color or insert their own image on each component and the background. For free drawings, players can use a free line to draw their name or picture as they want (Figure 11). Using this tool, users can create their own games.

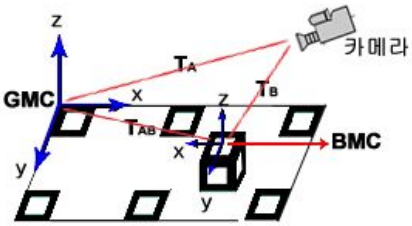
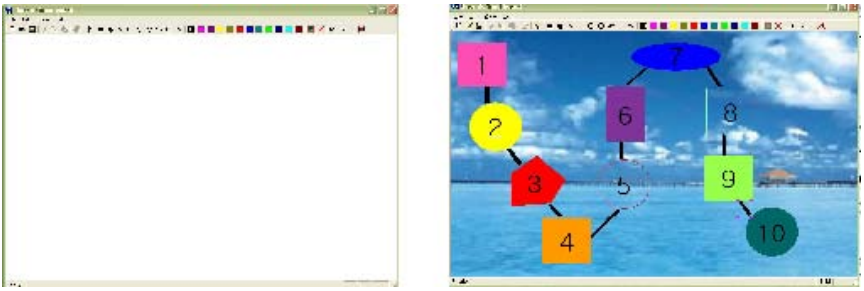


Fig. 10. Relationship between the board coordinate (GMC) and the operator marker or numbered marker coordinate (BMC)



(a) (b)

Fig. 11. Player defined board (a) Initial scene (b) An example

4 Conclusion

Game has been changed in variety of forms as it develops. Through this development, it shows that it can be applied to new purpose instead of being just a simple game. Within this, an educational game has been one of the categories that have been continually developed because it provides enjoyment and education.

This paper proposes an educational game in the Augmented Reality environment. In order to increase variety of computer games and learner's participation, AR is considered as a suitable environment. This is not limited to the proposed mathematical board game. If educational fields and AR technologies are brought together, learners can experience and learn while having fun; this can maximize the educational efficiency. The proposed mathematical educational game increases the enjoyment of learner and the educational effect, but the reliability requires improvement. The distance between camera and marker and various lighting condition cause unstable operations. Moreover, if the markers registered in the board are covered by the operator marker, the numbered marker or player marker, and can not be seen through the camera; virtual objects are not reliably overlaid. In order to use the system in real life, we have to solve these problems.

Acknowledgments

This work was sponsored and funded by Seoul R&BD Program.

References

[1] Lee, J.-b.: The present of serious game. Game industry journal, Korea, 10–14 (2007)
[2] Lee, S.-h., Jang, j.-w., Jun, c.-u.: The impact of serious games on the education of social problems. Game industry journal, Korea, 88–104 (2007)
[3] Son, B.-g., Go, B.-s., Gye, B.-g.: Strategy puzzle game: ALEPH, 1–21, 47–60 (2005)
[4] Woods, E., Billinghurst, M., Looser, J., Aldridge, G., Brown, D., Garrie, B., Nelles, C.: Augmenting the science centre and Museum Experience, Graphite, Singapore, 230–236 (2004)
[5] ZhiYing, S.Z., Cheok, A.D.: wlzQubesTM -A Novel Tangible Interface for Interactive Stroytelling in Mixed Reality. In: ISMAR, Japan, pp. 9–10 (2007)
[6] Lee, S.-H., Choi, J.-S.: AR Squash Game. In: ISMAR, Japan, pp. 4–8 (2007)
[7] Do, V.T., Lee, J.W.: Geometry Education using Augmented Reality. In: Mixed Reality Entertainment and Art Workshop2 in ISMAR, Japan, pp. 11–15 (2007)
[8] Jang, H.-b., Kim, J.-w., Lee, C.-w.: Augmented Reality Cooking System Using Tabletop Display Interface. In: International Symposium on Ubiquitous VR, Korea, pp. 36–37 (2007)
[9] http://ieeexplore.ieee.org/iel5/9382/29792/01357659.pdf?tp=&isnumber=&arnumber=1357659
[10] http://www.hitl.washington.edu/research/shared_space/download/
[11] http://www.powerpolitics.us/home.htm
[12] http://www.food-force.com
[13] http://www.nintendo.co.kr/www/soft/brain_01.php
[14] http://news.bbc.co.uk/1/hi/education/1879019.stm
[15] http://www.mxrcorp.com/
[16] http://www.hansmedia.com/
[17] http://www.euclideanspace.com/maths/geometry/rotations/euler/index.htm

Game-Based Learning Scenes Design for Individual User in the Ubiquitous Learning Environment

Stis Wu[1], Maiga Chang[2], and Jia-Sheng Heh[1]

[1] Dept. of Information and Computer Engineering, Chung-Yuan Christian Univ., Taiwan
stis@mcsl.ice.cycu.edu.tw, jsheh@ice.cycu.edu.tw
[2] School of Computing and Information Systems, Athabasca University, Canada
maiga@ms2.hinet.net

Abstract. A ubiquitous learning environment provides learners opportunities of observing and touching the learning objects around the learners according to their preferences and/or interests. Learners can still solve problems, answer questions, and propose their own questions, in the ubiquitous learning environment. When we plan to make learners study in the ubiquitous learning environment, there are some issues needed to be taken into considerations, for examples, learners' interests and preferences. How to encourage learners' interests is an important issue, this paper designs and builds learning scenes to offer learners the personalized learning services based on game concepts. Each learning scene may cover one or many learning spots, and each learning spot has different learning objects. We can construct a series of learning scenes dynamically for individual learner based on the learner's choices, preferences and interests. Furthermore, a learning path involves learning scene switch is also generated automatically for the learner. Several exhibition rooms and artifacts in a museum are used to demonstrate the idea and mechanism proposed by this research.

Keywords: game-based learning, scene, ubiquitous learning, pervasive learning.

1 Introduction

In recent year, ubiquitous learning extends e-learning from indoor to outdoor, moreover, not like most of mobile learning applications which usually only provides the knowledge of single domain in particular environment, the ubiquitous learning emphasizes to offer learners interdisciplinary knowledge domains in the real world [1][4]. Mobile learning provides both teachers and learners a new learning way in the e-learning field, however, there is still an unsolved research issue, which is the flexible learning issue. The learners' learning activities will be limited in the specific learning environment and/or the specific domain knowledge arranged in advance. Ubiquitous learning not only extends e-learning from indoor to outdoor but also extends mobile learning from specific learning environment and specific knowledge domain to anyplace and multi-discipline [5].

Ubiquitous learning also provides the learning activities which allow learners to observe and touch the learning objects in the real world based on learners' choices and preferences. If learners want, they still can do suchlike solving problems, answering

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 451–462, 2008.

questions, expressing and organizing questions, and even brainstorming, during their learning processes [2][3].

Some researchers have thoughts that there are four characteristics of games could enhance learning effects: (1) challenges; (2) fantasy; (3) curiosity; and, (4) control [6]. Regarding the first characteristic, the challenges, Malone has offered the instructional activities which have a various difficulty levels to different learners according to learners' abilities [7][8]. The different learners then will face different challenges.

Active learning makes learners learning better because they are doing what they feel interesting and they feel that they are controlling everything. Games can encourage learners to learn after classes. Learners are usually 'actively' to play their favorite games. So, it might be work if we apply game concepts, scenes and scenes switching, to help learners 'active' learn and have fun in the ubiquitous learning environment.

The research analyzes and constructs learning scenes to offer learners personalized game-based learning services in a ubiquitous learning environment, for examples, a museum and a zoo.

According to personalized context-aware knowledge structure, different learning scenes can be extracted for different learners. Through personalized scenes construction, the learning service can extract the learner's learning preferences and design the learning path for individual learner.

Section 2 introduces the related works of ubiquitous learning environment, game-based learning, different knowledge structures, and different learning strategies. The relations and definitions between the elements of the learning scene are described in Section 3. Section 4 uses a way to construct learning scenes and introduces how to switch between the scenes. In Section 5, a real example of ubiquitous learning in the museum with learning scenes is built to demonstrate the effects of this research. Finally, Section 6 makes a conclusion and discusses the possible future works about individual scenes layout and switching.

2 Scene Analysis

2.1 Scene Definitions

For storing knowledge and learning object information, this research defines "learning scene" to record all learning locations and objects in the ubiquitous learning environment. Besides, the scene is also used to gather the information about what learning objects exist in different learning locations.

A scene will cover many learning areas; each learning area is a physical place, for example, an exhibition room in a museum. In each learning area, there is at least one or more learning spot. A learning spot may cover one or many learning objects. Take a museum for example, there are many exhibition rooms (the learning areas). Each room has at least one learning spots. When the visitor stands at each learning spot, he/she can see one or many artifacts (the learning objects).

There are two major elements in a scene as Figure 1 shows: (1) Learning spot (spot), indicates a learning location which covered the learning objects in a personalized knowledge structure; (2) Learning object (object), denotes possible learning objects for the learning spot.

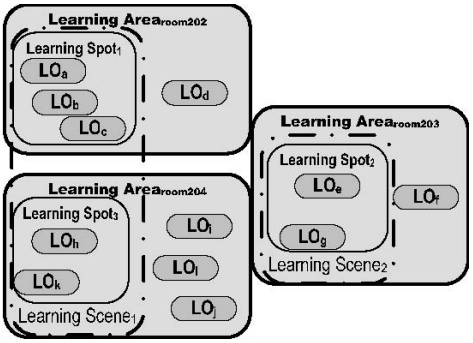


Fig. 1. Learining environment elements

2.2 Relations between Scenes and Personalized Context-Aware Knowledge Structure

The personalized context-aware knowledge structure then can be used to build learning scenes for learners. Because different learners with different preference will feel differently even when they are looking the same learning objects in the real world, hence, the suitable learning scenes to the learner can be built once the ubiquitous learning system has the personalized knowledge structure. Moreover, different learning activities can be given to learners depends on their preferences and learning objects' characteristics.

In Figure 2 there are two learners, Alex and Stis, have different viewpoints and/or preferences about the artifacts in the museum. The top part of Figure 2 shows Stis' personalized context-aware knowledge structure. Stis prefers the characteristic, "Dynasty", rather than "Color", hence, his personalized knowledge structure is constructed based on "Dynasty" characteristic. On the other hand, Alex has more interest in "Color", therefore his personalized knowledge structure root is constructed based on "Color" characteristic.

According to Wu's personalized knowledge structure [10], the scene subject for Stis is "Dynasty". The bottom part of Figure 2 shows how the ubiquitous learning system selects all learning objects related to the subject, "Dynasty", which Stis interests with to form a learning scene for Stis. Similarly, Alex's learning scene can be also constructed.

2.3 Game-Based Measurement with Learning Scenes

As the bottom part of Figure 2 shows, there are many learning objects matched the preference of a learner. The next problem is what learning object should the ubiquitous learning system suggests the learner learning and observing first.

Figure 3 shows an example of two learning scenes cover three learning spots and some learning objects. Assuming a learner stands at the learning spot2. At learning spot2, the learner can find two learning objects, {LO2, LO3}. Each learning object has its own characteristics about "Dynasty" subject. For example, the learning object LO2 may have characteristic suchlike "Items of daily use were specially produced for the Ming imperial court". Moreover, the degree of realizing these story-like characteristics

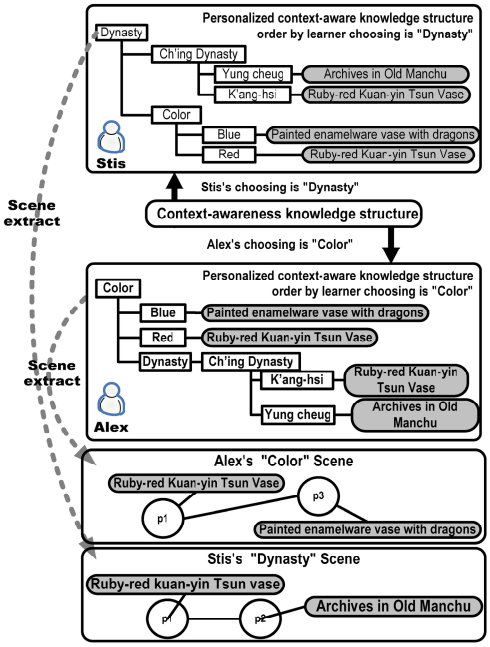


Fig. 2. Two personalized context-awareness knowledge structures

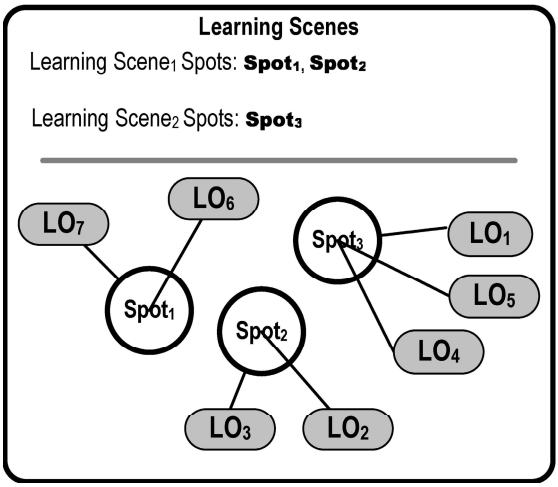


Fig. 3. Example of learning scene and objects

of learning objects can be seen as a quantitative measurement of learners' abilities and/or skill points based on the game-based learning theory.

This research defines the game-based measurement with the personalized learning scenes based on the probability theory. If a learner gets the idea from the questions, then the probability of specific characteristic will be raised and the probability

represent the learner's degree of mastery in the specific characteristic. When the overall probability of a learning scene achieved a threshold, the ubiquitous learning system can say that the learner clears the stage or levels up, and ready to challenge another stage or level (another learning scene).

3 Scenes Construction and Switching

3.1 Scenes Construction

This research develops a way to extract scenes from the personalized knowledge structure. The whole scene construction flow involves 5 phases as Figure 4 shows:

Phase I: Learning materials analysis, the learning objects in real environment have a lot of information (or also called characteristics), which might cover different domains or places. In this phase, we need to analyze the learning objects and its characteristics in the real world first. After this phase, the context-aware knowledge structure can be built.

Phase II: Basic personalized context-aware knowledge structure construction, because the personalized context-aware knowledge structure should be created according to the learners' preferences and interests. The basic personalized context-aware knowledge structure for individual learner can be constructed depends on which stories the learners interest with.

Phase III: Personalized context-aware knowledge structure refinement, even if two learners choose the same story in phase II, their interests still might be a little different. In order to precisely refine the personalized context-aware knowledge structure, the system asks the learner some advance questions about the characteristics of the learning objects which are involved in the story the learner has chosen. The personalized context knowledge structure is then refined according to the learner's feedback.

Phase IV: Personalized context-aware knowledge structure generation, after repeating the phase II and III, the correspondent learning objects and characteristics the learner might need and/or interest with are clear to the system. The learner's personalized context-aware knowledge structure is then can be generated in this phase.

Phase V: Learning scenes construction, the ubiquitous learning system can extract the learning objects that the learner may prefer according the personalized context-aware knowledge structure. Because the personalized knowledge structure can represent different knowledge domain and the learning objects locate at different locations, the ubiquitous learning system then construct the game-based learning scenes based on the distinction features of the selected learning objects, such as "location" and "domain", as Figure 4 shows.

3.2 Scenes Switching Principles

This research uses rough set theory [9] to develop scene switching methodology. In the rough set theory, data will be divided into three sets, including the Positive Set (POS), the Negative Set (NEG), and one or more Boundary Sets (BNDs). This research defines the positive set data as the learning objects which have more than one characteristic the

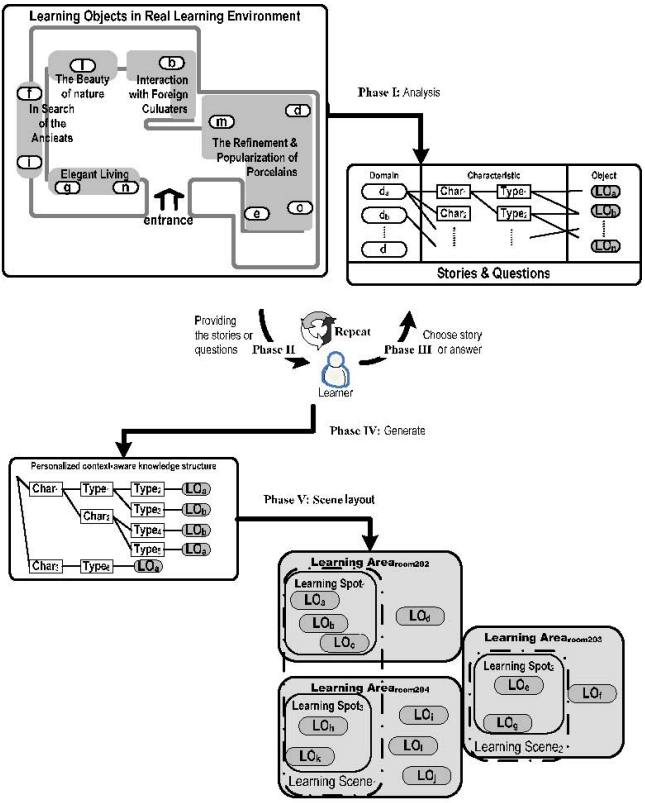


Fig. 4. The five phases to build individual scene

learner is interesting with, the learner has to observe the learning objects in the positive set. On the contrary, the learning objects in the negative set don't have any characteristic that the learner interest with. Unlike the positive set and the negative set, there are many boundary sets in the learning scene. The learning objects in a boundary set also have the characteristics that the learner interests with. The difference between the positive set and the boundary set is that all of the learning objects in the same boundary set have the same characteristics, and the learner only needs to observe one of the learning objects for the same boundary set when he/she does learning activities in the ubiquitous learning environment.

Figure 5 shows an example of using rough set to categorize learning objects involved in learning spots. At top part of Figure 5, there are two learning scenes cover three learning areas. Each learning scene has one or more learning spots, and each learning spot contains one or more learning objects. In the middle part of Figure 5, every learning object has many characteristics. The learning objects can be categorized into three sets. Table 1 lists a summary of the sets and its learning objects.

From Table 1 the ubiquitous learning system knows that the learner has to observe the learning object LOa. If the scene only includes four characteristics, including

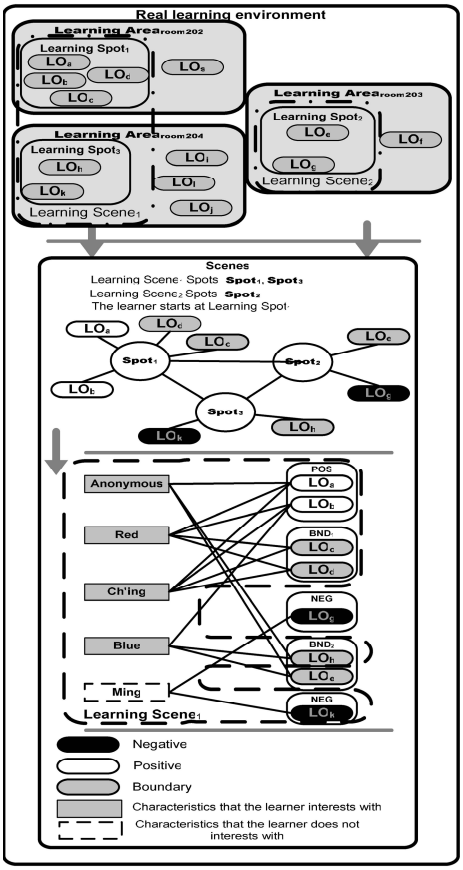


Fig. 5. The relation between characters and learning objects

Table 1. Three sets and its learning objects

POS	LO_a , LO_b
NEG	LO_g,LO_k
BND_1	LO_c, LO_d
BND_2	LO_e, LO_h

"Anonymous", "Red", "Ch'ing", and "Blue", then the ubiquitous learning system can do scene switching after the learner finished observing the LOa and LOb.

4 Complete Example

4.1 Game-Based Learning in Museum

By using the game concept to help learner learning in the ubiquitous learning environment, the first thing is to define the learner profile suchlike "ID", "ability", and

"major preference". The major preference includes the characteristics in the personalized context-aware knowledge structure suchlike "dynasty", "color", "function", and "author", in Figure 7. Ability is the learning object characteristics which learner has been observed. ID is the learner's name.

4.2 Scenario

This paper takes learning in a museum for example. In the museum, there are many artifacts with different subjects and/or topics exhibiting in the different rooms. For example, the Room 203 as Figure 6 shows, the "gray area" represents different subjects and/or topics (learning spots) suchlike "Elegant Living", "In Search of the Ancients", and "The Beauty of Nature"; "black circle" represents the artifacts (the learning objects) in the room; the number is the artifacts number. There are eleven learning objects in Room 203. This research considers a room as a learning area.

According the Phase I in Figure 4, Figure 7 shows the analysis results of learning objects. Figure 7 represents how to use context-aware knowledge structure to store a museum's learning objects and its characteristics.

The personalized context-aware knowledge structures and game-based learning scenes are constructed based on with the learners' choices and their answers about the questions of artifacts' characteristics as Figure 8 shows. The width or depth of personalize context-aware knowledge structure depends on how much the learners' interest with and how exquisite they know.

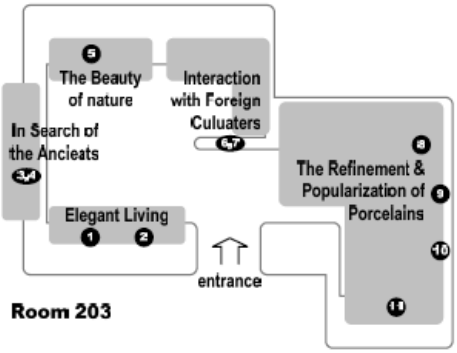


Fig. 6. Room 203 in a museum

Figure 9 shows the process of constructing the learning scenes after the learner answered the questions. After the learner answered the questions, the ubiquitous learning system can revise the personalized context-aware knowledge structure and find the preference learning objects and interesting characteristics from the personalized context-aware knowledge structure. Furthermore, using the personalized knowledge structure to realize the learning preference objects' locations, defining the learning spots, and constructing the learning scenes for individual learner as the bottom part of Figure 9 shows.

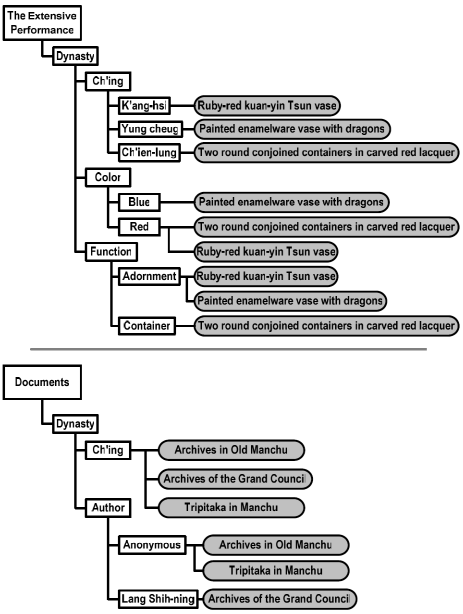


Fig. 7. The partial context-aware knowledge structure in a museum

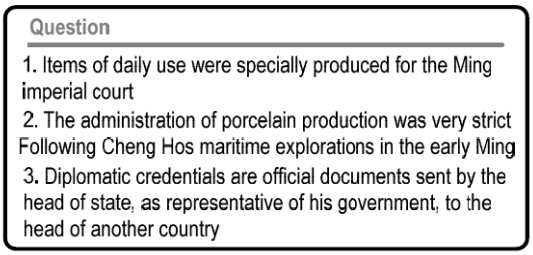


Fig. 8. The questions about "The Fashionable vs. the Antiquarian"

4.3 Individual Learning Path in Museum

With the learning scenes, this research uses the distances to plan the observation sequence for the learning objects in the positive set and boundary set that the learner may interest with.

The ubiquitous learning system then first picks the learning objects in the POS set, including "Ruby-red kuan-yin Tsun vase", "Two round conjoined containers in carved red lacquer", "Tripitaka in Manchu", "Archives of the Grand Council", and "Archives in Old Manchu". The learning path has to route around the POS set learning objects.

Figure 10 uses the "white ellipse" to represent the learning objects in the POS set, the must be observed learning objects. The scene can be switched from the Function Scene to the Document Scene after the learner has finished observing the white ellipse

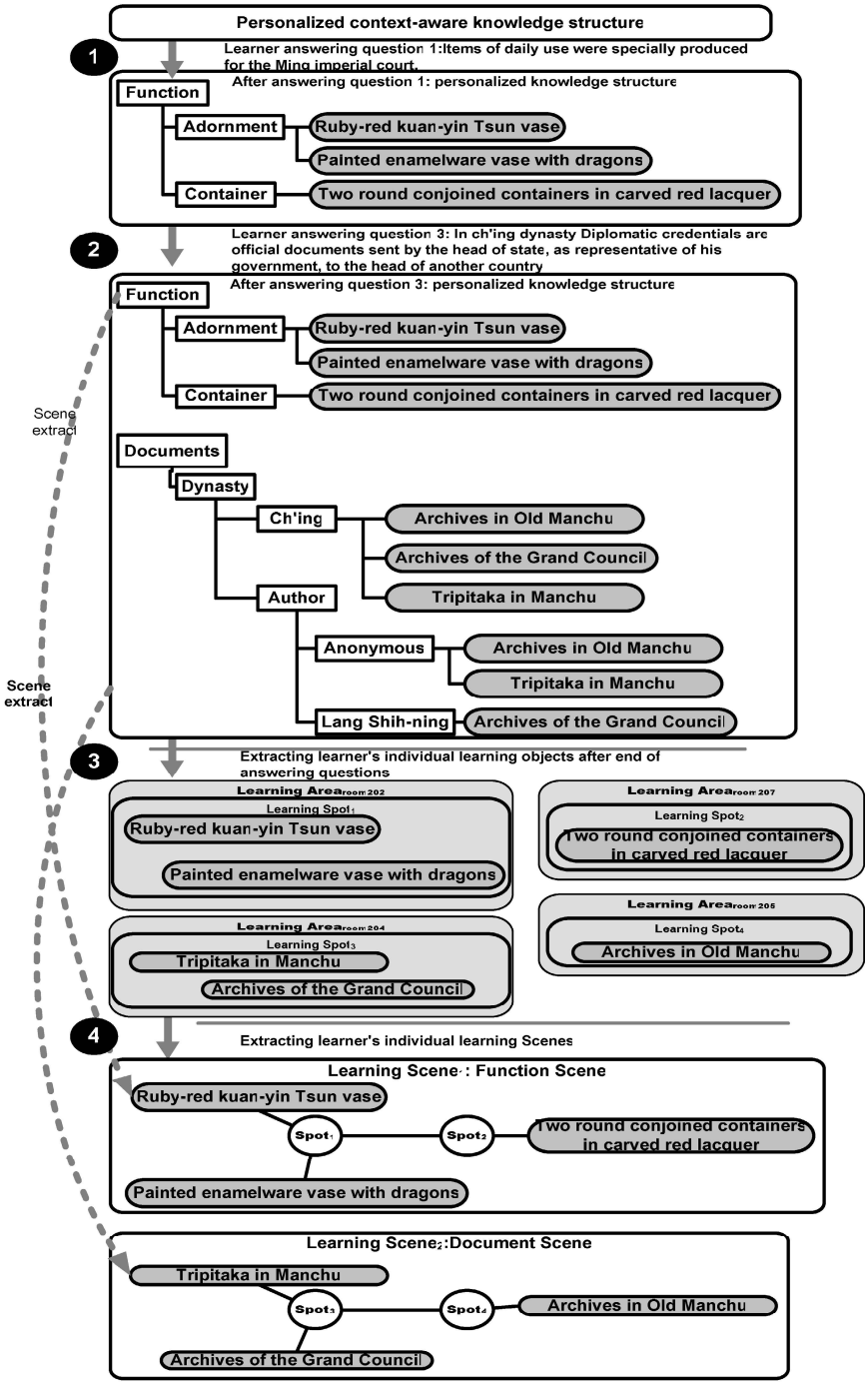


Fig. 9. Scenario constructing process

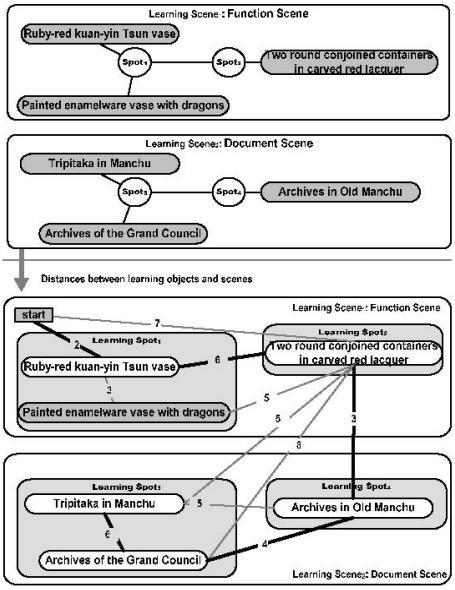


Fig. 10. Function scene and document scene relationship in scenario

learning objects. According to the distance, the ubiquitous learning system generates the learning guidance path: start Æ"Ruby-red kuan-yin Tsun vase"Æ"Two round conjoined containers in carved red lacquer"Æ"Archives in Old Manchu"Æ"Archives of the Grand Council"Æ"Tripitaka in Manchu".

5 Conclusions

There are still some research issues could be discussed and done in order to improve the scene switching and learning path generation. For example, as we mentioned, the challenge issue. Currently, the game-based ubiquitous learning system uses probability as the measurement of learner (player) abilities. The next step is how to provide an automatically generated and non-interrupt adaptive test for such kind of ubiquitous learning environment. Regarding the learning path generation issue, the current research and system uses only distance between the learning objects. As we all know, it is just the simplest solving way, in the future research, other factors suchlike learning scenes should be taken into consideration.

References

[1] Chang, A., Chang, M.: Adaptive Mobile Navigation Paths for Elementary School Students' Remedial Studying. In: Proceedings of the IEEE International Conference on Interactive Computer Aided Learning (ICL 2006), Villach, Austria, September 27-29 (2006)

[2] El-Bishouty, M.M., Ogata, H., Yano, Y.: Personalized Knowledge Awareness Map in Computer Supported Ubiquitous Learning. In: Proceedings of the 6th IEEE International Conference on Advanced Learning Technologies (ICALT 2006), Kerkrade, The Netherlands, July 2006, pp. 817–821 (2006)

[3] Felder, R.M., Brent, R.: Learning by Doing. Chemical Engineering Education 37(4), 282–283 (2003), Retrieved on January 15, 2008, from `http://www4.ncsu.edu/unity/lockers/users/f/felder/public/Columns/Active.pdf`

[4] Hwang, G.-J.: Criteria and Strategies of Ubiquitous Learning. In: Proceedings of the IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC 2006), Taichung, Taiwan, June 2006, pp. 72–77 (2006)

[5] Hall, T., Bannon, L.: Designing ubiquitous computing to enhance children's learning in museums. Journal of Computer Assisted Learning 22(4), 231–243 (2006)

[6] Lepper, M.R., Malone, T.W.: Intrinsic motivation and instructional effectiveness in computer-based education. Aptitude, Learning, and Instruction, III: Conative and Affective Process Analysis, pp. 255–286. Erlbaum, Hillsdale, New Jersey (1987)

[7] Malone, T.W.: Toward a theory of intrinsically motivating instruction. Cognitive Science 5(4), 333–369 (1981)

[8] Malone, T.W., Lepper, M.R.: Making learning fun: A taxonomy of intrinsic motivations for learning. Aptitude, Learning, and Instruction, III: Conative and Affective Process Analysis, pp. 223–253. Lawrence Erlbaum Associates, Hillsdale (1987)

[9] Pawlak, Z.: Rough Sets: Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers, Norwell, MA

[10] Wu, S., Chang, A., Chang, M., Liu, T.-C., Heh, J.-S.: Identifying Personalized Context-aware Knowledge Structure for Individual User in Ubiquitous Learning Environment. In: Proceedings of the 5th IEEE International Conference on Wireless, Mobile, and Ubiquitous Technologies in Education (WMUTE 2008), Beijing, China, March 23-26, 2008 (in press, 2008)

Learning Models for the Integration of Adaptive Educational Games in Virtual Learning Environments

Javier Torrente, Pablo Moreno-Ger, and Baltasar Fernandez-Manjon

Dpto.Ingeniería del Software e Inteligencia Artificial.
Facultad de Informática, Universidad Complutense de Madrid
28040, Madrid. Spain
{jtorrente, pablom ,balta}@fdi.ucm.es

Abstract. There is a trend in Virtual Learning Environments (VLE) towards flexible and adapted learning experiences that modify their contents and behavior to suit the needs of different learners. On the other hand, the use of educational videogames is also an emerging trend to address diverse aspects, such as student engagement and exploratory learning. Additionally, videogames support adaptation in a natural way. This suggests that they may be a good vehicle to enhance the adaptive features of VLE. However, all those ideas are partially disconnected and, in spite of all the work done, there is still a need for effective learning models that leverage the potential of games, integrating them with the available learning materials and VLE. In this work, we discuss such models and describe how the <e-Adventure> educational game platform supports them.

Keywords: educational games, learning models, user adaptation.

1 Introduction

Nowadays, adaptation to the needs of different learners and contexts is becoming an increasingly important aspect of Virtual Learning Environments (VLE) [1, 2]. This is a result of the need to reach users anywhere and anytime combined with the flexibility of web technologies. Typical adaptation mechanisms build student profiles based on learner preferences, portfolio, previous knowledge, educational objectives, and, in some cases, even different learning styles [3, 4].

Another increasingly important aspect in the field of educational technologies is the inclusion of digital games in educational environments. Works such as [5, 6] identify the importance of motivation in the learning process and, from that base, other authors discuss the benefits of using videogames to enhance the quality of the learning experiences [7, 8]. Moreover, other authors defend that the characteristics that make videogames attractive (immersion, short feedback cycles, scaffolded learning, perception of progress, etc.) are also key elements in any effective learning process [9, 10].

Additionally, the characteristics of game-based learning suggest its potential benefits when applied to adaptive online learning. Adaptation is a pervasive feature of commercial videogames, since they are practically required to support different difficulty levels in order to cater to the broadest possible audience. Moreover, unlike with

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 463–474, 2008.

an HTML or PDF document, in a game we can monitor very closely the interaction between the student and the game and use this information as part of our adaptation cycles [11].

However, in spite of all the work created so far, there still is a need for effective learning models that leverage the potential of games integrating them with the available learning materials and VLE. The use of educational videogames as content in adaptive Virtual Learning Environments in a sensible way is not a trivial problem. To begin with, we cannot force all the broad variety of students to learn from game-based solutions. Some students may lack the proper equipment, some public computers may restrict the types of content that can be accessed, and some students may simply refuse to play the games seeing them as a waste of their time [12]. Even from an adaptation perspective, some student profiles may require a very closely guided learning experience as opposed to the exploratory freedom offered by most games. In such cases, a conventional approach based on HTML, PDF and multimedia content could be more appropriate. Finally, traditional web content is more accessible for people with special accessibility needs, for example using text-to-speech tools for blind people.

For these reasons, in the context of adaptive online learning, the employment of educational games should not replace other approaches, but try to complement the learning experience by integrating both alternatives in richer learning environments. We propose a new model of integration, in which game-based content is blended and complemented with traditional web contents as opposed to substituting them.

However, in order to apply such a model, it is necessary to address a number of technical issues. In this work we describe how the <e-Adventure> educational game platform [13]can be used to support this approach, facilitating the development of educational adventure games and their integration with Virtual Learning Environments.

This work is structured as follows: section 2 describes some of the issues regarding adaptation and game-based learning that supply the base for this work. Then, on section 3, we discuss the issues that emerge when substituting traditional content with game-based content and propose a learning model that mixes both. On section 4, we describe our <e-Adventure> platform, which can be used to implement that learning model. Finally, some conclusions and future work are discussed in section 5.

2 Game-Based Learning and Adaptive Learning

2.1 Adaptive Online Learning

When we speak about adaptation in a Virtual Learning Environment, we are usually speaking about a system that gathers information about its students and then uses the information stored in the students' profiles to customize the content delivered to learners and/or the activities they must perform [3]. Therefore, adaptation usually deals with two different problems: Gathering information about the student and then modifying the learning experience. The adaptation may be addressing a wide variety of aspects, such as customizing the Graphical User Interface (GUI), supporting different learning objectives or initial levels of knowledge, adaptation to different use contexts (e.g. a public computer vs. a desktop computer at home), and even supporting

different learning styles. For the scope of this work, we will only focus on the technical needs when it comes to adaptation in terms of a) different levels of initial knowledge; and b) different learning styles.

Adaptation to different levels of initial knowledge requires finding out the current level of each student in order to adapt the content accordingly. For instance, it could be addressed by filtering basic content for those students with a certain level of knowledge, so they could effectively focus their effort

Regarding the adaptation to different learning styles, we are aware that this is still a controversial field. In spite of long empirical efforts to pin them down, the identification of learning styles remains elusive [14]. However, most people with teaching experience acknowledge intuitively that there are differences in how their students learn [15]. For this work we will assume a very rough classification of student profiles, cataloguing them according to whether they are able to self-regulate their own learning processes (preferring a free and exploratory approach), and those who prefer close teacher control and guidance.

2.2 Adaptive Game-Based Learning

The advantages of integrating games in educational environments have been widely discussed in the literature [16-18], However, we would like to point out the properties that make them a particularly adequate medium for adaptive learning.

The entertainment videogame industry has grown and now it is a mature industry that caters for all ages and genders. Driven by a commercial pressure to entertain different player profiles, successful games have developed sophisticated adaptation mechanisms. Most games adapt their behavior to suit different levels of proficiency, adjusting the difficulty of the game (sometimes even automatically [19]). Some games even adapt to suit different playing styles so that each player enjoys the game experience as much as possible. It can be concluded that current game technology inherently supports the features that an adaptive learning experience demands. Additionally, their high interactivity and complexity mean that the content can be adapted both in general behavior and in fine-grained details. Moreover, games can be designed and implemented with the means to track the progress and actions of learners while they are playing, gathering valuable information for both adaptation and assessment purposes.

Therefore, a game can implement a complete adaptive learning cycle, both by gathering information about the player during the game and by modifying its behavior as needed. The game pattern shown in Figure 1 is an example of an adaptive game-based learning architecture. It contemplates different game itineraries for different students. On each itinerary the game behaves differently. The game can exhibit different behaviors in order to support different learning styles (for example, giving the player more or less freedom to explore) or different learning objectives (for example, omitting some advanced details from the game). Additionally, the games can skip those levels that are too basic for the student's initial level of knowledge.

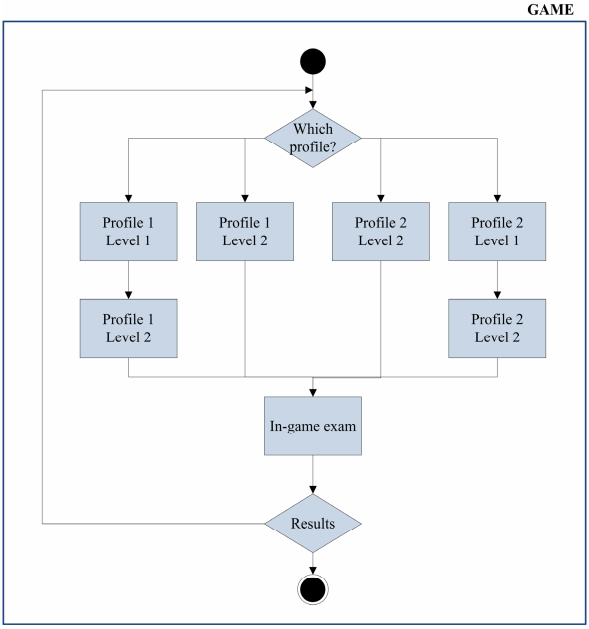


Fig. 1. A game pattern for adaptive game-based learning. The game can exhibit different behaviors to support different learning styles (Profile 1, 2…) and can omit certain levels that are too basic for the level of initial knowledge. The game also gathers information from the interaction during the in-game exam and uses it to modify its own behavior.

This pattern also contemplates using the game mechanics for assessment purposes, monitoring the activity of the student during an in-game exam. Whilst the game is played, a lot of interaction between the learner and the game is produced. From the monitorization of all these interactions it is possible to infer data about the learners that could be used to categorize them into one of the learning styles previously defined. In some cases, if the results achieved by the student are insufficient, it is also possible to reassess the profile of the student and run the game again with a different profile.

3 Combining Educational Videogames and Traditional Content in Adaptive Learning Patterns

Most game technologies and genres can easily support the game pattern outlined in the previous section. However, it must be done cautiously. Game-based learning may not be appropriate for all the students all the time or even for all the possible subjects. Thus, the integration of these games into the Virtual Learning Environments should go beyond simply deploying educational videogames instead of web content.

In this section we describe the issues identified when it comes to integrating adaptive games in Virtual Learning Environments, and then we propose a model supporting such integration.

3.1 Educational Games: User Interaction and Access-Related Issues

As it has been stated in the previous section, educational videogames can be a vehicle for the introduction of complex adaptation procedures in the learning experience, enhancing motivation and providing an immersive domain to interact with. However, game-based approaches would neither suit every student's tastes nor be adequate in all contexts and situations.

First, the technological complexity of videogames is an issue. Most games demand high system requirements and some students may find their computers unable to execute them. Similarly, as videogames are complex pieces of software, its use is restricted in most public and private systems beyond the personal field. Additionally, there is an emerging trend towards the use of mobile devices (e.g. PDAs, mobile phones, etc.), in what has been named m-Learning (mobile learning) [20] and mobile learners usually will favour traditional web-based content.

Apart from the technological issues, sometimes the students themselves would decide not to use these games. Games are usually time-consuming, as they require getting familiarized with the environment (i.e. the domain of study) and sometimes even learning how to use and interact with the universe devised in the game. Moreover, some students are averse to videogames. A person with no experience playing commercial videogames would find in educational videogames an extra and superfluous challenge to waste their efforts on instead of an additional motivation, as witnessed and reported in [12]. If addressing adaptation in terms of students' preferences and learning styles is desirable, we should avoid forcing those students who distrust videogames to learn from game-based contents.

Another issue would be students that cannot interact with game-based contents. For instance, blind people would find it impossible to learn through videogames, as visual interaction cannot be replaced in games with a further use of the other senses. In these cases, alternative learning itineraries, like HTML web pages that can be interpreted by a text-to-speech tool, should be provided, following the current research aiming to provide access for all to Information Systems [21].

Finally, game-based solutions are not recommended for those students that feel overwhelmed by the freedom of exploration provided in games (even when such freedom can be gauged and adapted transparently in the games). In those cases traditional content, in which the interaction during the learning experience is more rigid seems to be more suitable.

3.2 Integration of Games with Traditional Content

Let's consider a typical scenario: a learning module composed by a number of Learning Objects (e.g. HTML documents), which is already deployed in a VLE and being accessed by students via a web browser. However, the instructors decide to seek alternative didactic methodologies including adaptation and educational games. An adaptive game-based version of the content is designed to suit a profile of students who probably have no study habit but have game habit. While carrying out the integration of the new game-based content into the learning module, the instructors find two main issues that need to be tackled:

On the one hand, the educational game is more costly to develop and it may be complicated to ensure that guarantees the accomplishment of the same learning objectives. This adds a new burden for the instructor or the person entrusted to produce and maintain the learning content. On the other hand, the inclusion of the game in the learning module would require reshaping the content and the pedagogical approach in order to fit in a game-based delivery.

These issues could be addressed by supplying instructors (or content designers) with mechanisms for the automatic integration of the existing content into the new game-based content so that it can be easily accessed from the game, with no extra effort for the instructors. In this way, both learning itineraries (traditional and game-based) would incorporate the same information.

3.3 Description of the Resulting Adaptive Learning Pattern

Taking into account the aforementioned aspects, we propose an adaptive model (Figure 2) with two adaptation layers. Firstly, the VLE decides whether a game-based or a traditional HTML approach is more appropriate for the learner according to the profile. This decision may depend on the requirements of the student (learning styles, disabilities, etc.), the current context (a mobile device or a short consultation session) or the student's preferences. Secondly, when game-based content is chosen, a more fine-grained adaptive mechanism is applied to adjust the game's behaviour in the terms discussed thus far.

These adaptive mechanisms are supported mainly by a test performed at the beginning of the learning experience. This pre-test, checks whether the game-content is suitable for the students and their initial level of knowledge. In some cases, the student profile kept by the VLE will be sufficient to decide the shape of the content to be delivered, making this test unnecessary. In other circumstances (e.g. when the student is a novice using the system) a questionnaire could be given to the student in order to find out their knowledge, preferences and perhaps some information about their "learning taste", so that their learning style can be inferred. The complexity of these tests is in accordance with the adaptation mechanisms that we wish to apply. Most models that attempt to capture learning styles include methodologies to infer the learning styles of each student (like Vermunt's Inventory of Learning Styles questionnaire [22]).

The adaptation within the game can fit different learning styles by displaying different game behaviours (such as biasing the behaviour towards guided or exploratory styles) and also support different levels of initial knowledge by skipping those sections which appear too simple for a particular student.

After the learning experience has been completed, a second test (post-test) is performed. The results of this test (that can be obtained within or outside the game) can be used to refine the student's profile in order to improve future adaptation decisions. If a game-based content was the choice, an in-game exam could be the source of the assessment, along with other data gathered through the monitorization of the interaction of the student with the game. On the other hand, if conventional content was chosen, a traditional online exam would provide the information.

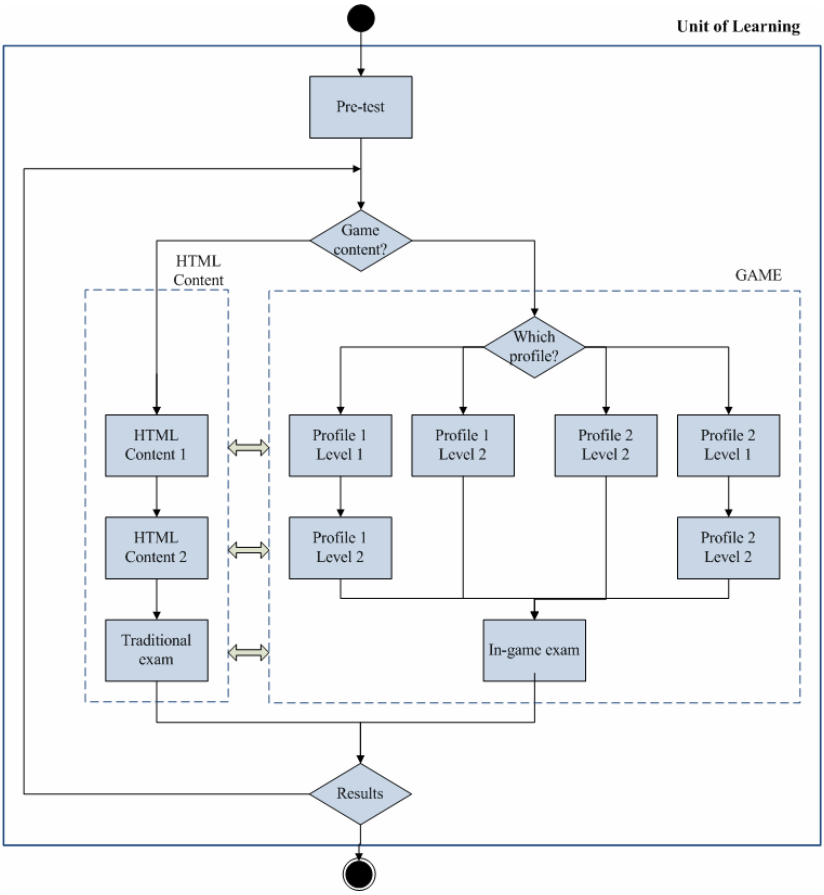


Fig. 2. Adaptation model considering two different learning itineraries: HTML-based and GAME-based. The adaptation is focused on two conditions: learning styles (different content paths, different profiles in the game) and prior-knowledge (initial levels of the game can be skipped).

Regarding the issues when integrating game-based content in the learning experiences, we consider that both learning itineraries should not be disconnected. As shown in figure 2, the itineraries are linked allowing the games to access the HTML content as suggested in section 3.2.

4 Implementation of the Adaptive Pattern Using <e-Adventure>

<e-Adventure> is an educational game engine designed to facilitate the creation of interactive educational content, focusing on pedagogical aspects such as adaptation and assessment [11]. It is a complete authoring environment for graphical *point and click* adventure games, built around an XML-storyboard, which supports the adaptive model proposed in this article.

The platform can be integrated with Virtual Learning Environments. When deployed from a standards-compliant VLE, the implementation of the engine can query the LMS for a set of properties that are used to adapt the game. The games are defined so that the different values of those properties will change the initial state of the game and consequently the game will be adapted, thus supporting the second adaptation level of the general pattern described in section 3. The adaptive cycle is closed by in-game tests, which are automatically assessed by the engine providing the necessary feedback to readjust the student's profile if necessary.

Another issue to address is how to effectively provide domain information to learners while playing because this aspect is not usually present in commercial games. Following the traits of the adventure game-genre [23], the interactive conversations with the characters that the players encounter inside the game represent a good source of information. However, interactive conversations with other characters are not always ideal. Some contents may require being delivered through alternative metaphors. When a large amount of data has to be delivered, conversations are not the most "natural" channel. Long conversations will prompt the students to loose their attention and focus. Moreover, embedding large amounts of information in conversations reduces its availability.

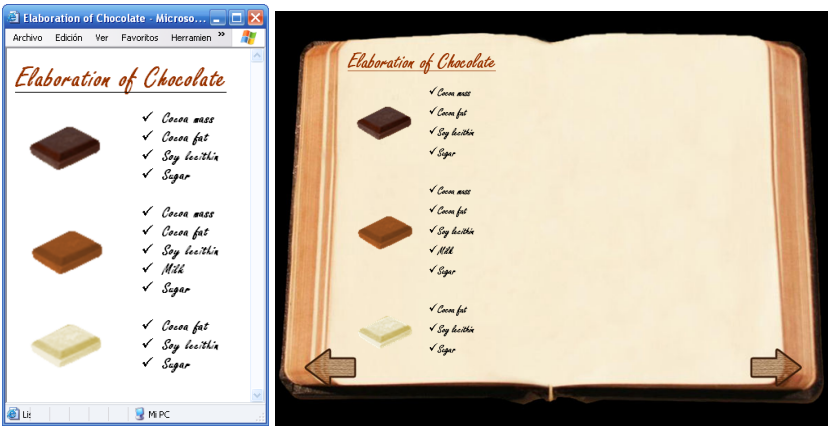
It is often desirable to make sure that some reference information is at the reach of the student at every moment. A first solution would be to allow the students to consult separate online materials containing the information, although this can be cumbersome for the student who is forced to switch back and forth between the game and the contents. For these reasons, <e-Adventure> includes the notion of *in-game books*. Those books are available to the learner at any moment to be looked up, supporting both text and images.

The books can be specified in <e-Adventure> by two different approaches: On the one hand, the content of the books can be marked up along with the XML-storyboard, as depicted in Figure 3. On the other hand, the definition of the contents of the books can be detached from the game content by referring to a web page. In this manner, the

```
<book id="BookElaborationChocolate">
  <resources>
    <asset type="background" uri="assets/background/book.jpg"/>
  </resources>
  <text>Elaboration of chocolate
    <img src="assets/image/DarkChocolate.png"/>
          Cocoa mass
          Cocoa fat
          Soy lecithin
          Sugar
          --
    <img src="assets/image/MilkChocolate.png"/>
          Cocoa mass
          Cocoa fat
          Soy lecithin
          Milk
          Sugar
          --
    <img src="assets/image/WhiteChocolate.png"/>
          Cocoa fat
          Soy lecithin
          Sugar
          --
  </text>
</book>
```

Fig. 3. A fragment of a marked-up in-game book in <e-Adventure>. The figure depicts how a book is represented in the storyboard (left) and how it is visualized in the game (right), as an actual book.

content of the book is retrieved from a URL where an XHTML document is located and subsequently displayed in the game, as depicted in figure 4. A first advantage of detaching the content of the books from the storyboard (where books are marked up) is that the production and maintenance of book content is eased as the instructor can edit and organize the information using HTML authoring tools. However, the main advantage of this approach is how it supports the integration of existing web content in <e-Adventure>. This can be leveraged to support the key issues from the adaptive learning model described in section 3, allowing direct access to the web-based content deployed on the VLE from inside the game.




```
<book id="BookElaborationChocolate">
   <resources>
         <asset type="content" uri="assets/books/elaboration/content.htm"/>
         <asset type="background" uri="assets/background/book.jpg"/>
   </resources>
</book>
```

Fig. 4. Reference information can be designed using web authoring tools as an XHTML document (left). Then, <e-Adventure> renders this document into a better designed book (right). The piece of XML (below) shows how the book is defined by just referring to its assets (background image and content web page).

5 Conclusions

In this work, we have argued the necessity of bringing together two increasing trends in online education: adaptation (in order to suit the broader range of people, reaching users anywhere and anytime) and the application of videogames to educational purposes. Nonetheless, it has been remarked that traditional content should not be totally replaced, as game-based solutions are not always the best choice for everyone at anytime. Instead of that, both approaches should coexist in the Virtual Learning Environments, combining the advantages of both approaches. In our opinion, the achievement of the aforementioned goals will need the research to move towards the development of learning models that integrate in a sensible way both key concepts (adaptation and use of educational videogames).

Considering all these reasons, we have outlined an adaptive learning model supporting a full adaptation cycle and the integration of game-based content with traditional alternatives. This model addresses adaptation in two layers: First, we introduce the capacity to diversify the learning experience, supporting different itineraries (including game-based and traditional content) in order to suit the broadest range of learners and situations. The second adaptation stage takes place when game-content is chosen, leveraging the characteristics of videogames to provide a much more fine-grained adaptation mechanism. The first adaptation step is thus to decide what sort of content is to be delivered according to the learner's profile. Therefore, this layer not only represents a first adaptation step, but also a general guideline for the integration of educational videogames in existing VLEs: Traditional content should be kept and offered as an alternative instead of replaced.

Another relevant concept discussed is the relevance of integrating the existing web content inside the games. Some types of information do not translate easily into game features, and simply work better in their textual form.. A first idea would be to allow the students to consult the online materials while playing, although this can be cumbersome for the student who is forced to switch back and forth between the game and the contents. Another approach would be to embed all the HTML content into the game. This allows the student to use the reference materials from within the game. However, duplicating the content is not a sustainable approach from a content adaptation and maintenance perspective. The solution presented here is to link these contents from the games, having the game engine render web content retrieved directly from the VLE.

It is important to remark that the proposed model is a simplification devised to display the features that should be supported by the technology employed to perform the integration of adaptive games in online education. For the sake of simplicity, the model only displays adaptation driven in terms of learning styles and prior-knowledge, although it is broad enough to support other adaptation approaches.

As described in section 4, this model can be implemented using the <e-Adventure> educational game platform. The assessment and adaptation mechanisms offered by the platform support the adaptive features described in the adaptive game model outlined in section 2.2. Additionally, since <e-Adventure> can be deployed in standards-compliant Virtual Learning Environments, it can support the general adaptive cycle described in section 3.3, facilitating the proposal of alternative game-based itineraries without detaching the game experience from the rest of the learning process. Finally, the capacity of the platform to render HTML-based Learning Objects inside the game using a book metaphor solves the issues related to the integration of existing web content inside the games in a sensible way.

As a final remark, it must be mentioned that the use of in-game books to provide the learner with reference information, may be a double-edged weapon: an abuse of in-game books is prompt to make the learning experience boring and give the narration a slow pace, instead of making it more motivating and attractive. When talking about educational videogames we should always bear in mind that both *fun* and *educational content* should be present and balanced; otherwise we will have the "benefits" of a bad learning experience (sometimes referred as "eduboring") but at a higher cost. We are trying to achieve such balances developing learning modules that

implement this model in the field of Computer Science teaching. The results obtained will be useful to gauge a further detailed model and propose future lines of work.

Acknowledgements. The Spanish Committee of Science and Technology (projects TIN2005-08788-C04-01, FIT-360000-2007-23 and TIN2007-68125-C02-01) has partially supported this work, as well as the Regional Government of Madrid (grant 4155/2005) and the Complutense University of Madrid (research group 921340).

References

[1] Burgos, D., Tattersall, C., Koper, E.J.R.: Representing adaptive and adaptable Units of Learning. How to model personalized eLearning in IMS Learning Design. In: Fernández Manjon, B., et al. (eds.) Computers and Education: E-learning - from Theory to Practice. Springer, Heidelberg (2007)

[2] Sancho, P., Martínez-Ortiz, I., Fernández-Manjón, B., Moreno-Ger, P.: Development of a Personalized e-Learning Experience Based on IMS Standard Technologies. In: Mendes, A.J., Pereira, I., Costa, R. (eds.) Computers and Education: Towards Educational Change and Innovation, pp. 73–82. Springer, London (2008)

[3] Brusilovsky, P.: Adaptive Educational Systems on the World-Wide-Web: A Review of Available Technologies. In: WWW-Based Tutoring Workshop at 4th International Conference on Intelligent Tutoring Systems (ITS 1998), San Antonio (1998)

[4] Hannafin, M.J., Land, S.M.: Student centered learning and interactive multimedia: status, issued, and implication. Contemporary Eduaction 68(2), 94–99 (1997)

[5] Malone, T.W., Lepper, M.R.: Making learning fun: A taxonomy of intrinsic motivations for learning. In: Malone, T.W., Lepper, M.R. (eds.) Aptitude, learning and instruction III: Cognitive and affective process analysis, pp. 223–253. Lawrence Erlbaum, Hillsdale (1987)

[6] Lepper, M.R., Cordova, D.I.: A desire to be taught: Instructional Consequences of Intrinsic Motivation. Motivation and Emotion 16, 187–208 (1992)

[7] Garris, R., Ahlers, R., Driskell, J.E.: Games, Motivation and Learning: A Research and Practice Model. Simulation & Gaming 33(4), 441–467 (2002)

[8] Jenkins, H., Klopfer, E., Squire, K., Tan, P.: Entering the Education Arcade. ACM Computers in Entertainment 1(1) (2003)

[9] Gee, J.P.: What video games have to teach us about learning and literacy, p. 225. Palgrave Macmillan, New York, Basingstoke (2003)

[10] Van Eck, R.: Building Artificially Intelligent Learning Games. In: Gibson, D., Aldrich, C., Prensky, M. (eds.) Games and Simulations in Online Learning: Research and Development Frameworks. Information Science Publishing, Hershey (2007)

[11] Martinez-Ortiz, I., Moreno-Ger, P., Sierra, J.L., Fernández-Manjón, B.: Production and Deployment of Educational Videogames as Assessable Learning Objects. In: First European Conference on Technology Enhanced Learning (ECTEL 2006). LNCS. Springer, Heidelberg (2006)

[12] Squire, K.: Changing the game: What happens when video games enter the classroom. Innovate, Journal of Online Education, 1(6) (2005)

[13] Moreno-Ger, P., Martínez-Ortiz, I., Sierra, J.L., Fernández-Manjón, B.: A Content-Centric Development Process Model. IEEE Computer 41(3), 24–30 (2008)

[14] Mayes, T., De Freitas, S.: Review of e-learning theories, frameworks and models (2004)

[15] Coffield, F., Moseley, D., Hall, E., Ecclestone, K.: Learning styles and pedagogy in post-16 learning, Learning Skills Research Centre (2004)
[16] Aldrich, C.: Learning by Doing: A Comprehensive Guide to Simulations, Computer Games, and Pedagogy in e-Learning and Other Educational Experiences. Pfeiffer, San Francisco (2005)
[17] Kirriemur, J., McFarlane, A.: Literature review in games and learning., NESTA Futurelab, 8 (2004)
[18] Mitchell, A., Savill-Smith, C.: The Use of Computer and Videogames for Learning: A Review of the Literature. Learning and Skills Development Agency, Trowbridge, Wiltshire (2004)
[19] Hunicke, R., Chapman, V.: AI for Dynamic Difficulty Adjustment in Games. In: 19th Nineteenth National Conference on Artificial Intelligence (AAAI 2004), AAAI Press, San Jose, California (2004)
[20] Savill-Smith, C.: The use of palmtop computers for learning: a review of the literature. British Journal of Educational Technology 36(3), 567–568 (2005)
[21] IMS Global Consortium. IMS Guidelines for Developing Accessible Learning Applications, Version 1.0 White Paper (2005) (cited March 2008), http://www.imsglobal.org/accessibility/index.html
[22] Vermunt, J.: The regulation of constructive learning processes. British Journal of Educational Psychology 68(2), 149–171 (1998)
[23] Amory, A.: Building an Educational Adventure Game: Theory, Design and Lessons. Journal of Interactive Learning Research 12(2/3), 249–263 (2001)

The Potential of Interactive Digital Storytelling for the Creation of Educational Computer Games

Sebastian A. Weiß and Wolfgang Müller

PH Weingarten, University of Education
Media Education and Visualization Group (MEVis)
Kirchplatz 2, 88250 Weingarten, Germany
{weiss, muellerw}@ph-weingarten.de
http://www.ph-weingarten.de

Abstract. The usage of computer games for educational purposes is currently widely discussed. But, to what extend do computer games cope with the requirements for learning? In this paper we discuss the application of digital games to learning and the problems and demands resulting from this integration. The question on how to integrate learning successfully with elements of play and games seems to be unsolved. We propose to integrate Interactive Digital Storytelling (IDS) with game-based learning (GBL) as a concept for building educational computer games. Furthermore, we discuss the requirements for effective learning applications based on the IDS paradigm.

Keywords: Interactive Digital Storytelling, Game-based Learning, Serious Games, Educational Computer Games, Scenejo, Killer Phrase Game, Learning.

1 Introduction

E-learning has been considered one of the most promising new markets during the last years. Major activities covered the development of learning environments and standards for the exchange of learning material. From the point of view of learning concepts, the resulting material adapted mostly ideas from behaviorism, resulting in a more or less rigid structure of electronic learning environments, making it hard to provide self-guidance to a learner. On the other hand, e-learning offers the chance to deliberate about learning concepts in a new way.

Game-based learning (GBL) describes the application of computer games for learning, or, as Prensky has put it, it "is precisely about fun and engagement, and the coming together of serious learning and interactive entertainment into a newly emerging and highly exciting medium" [1]. Furthermore, Gee [2] has analyzed that computer games are new media that let children and adults experience a learning effect while enjoying themselves.

Learning with Computer Games

Play has often been proposed as an effective learning paradigm (see for instance [3], [4], and [5]), and computer games have been attributed to further intrinsic motivation

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 475–486, 2008.

by elements such as challenge, fantasy, curiosity, and choice and control [6]. Moreover, achieving a flow experience [7] may also provide further extrinsic motivation [8]. Furthermore, narrative in games may provide a cognitive framework for problem-solving [9]. However, consolidated findings on the effectiveness of game-based learning are limited [10].

The fact that due to the high cost for developing good and complex computer games, educational games do often not meet necessary quality standards in terms of technology, complexity of the scenario, story, and didactic contents (see for instance [11], [12]), represents a major problem in this context. As a consequence, commercial of-the-shelf (COTS) computer games are frequently being applied to teaching and learning. However, commercial games are in general not designed for applications in learning, and content may be inaccurate or incomplete. Applying COTS games to learning requires therefore careful analysis of the game's strengths and weaknesses as well as a matching of the game's content to learning goals [13].

A number of commercial learning games exists targeted for specific learning scenarios. Yet, most of these follow behaviouristic approaches, (e.g., *Mathematikus* [14], *Physicus*, *Informaticus* [15]). As such, they do not reflect today's understanding on how to design successful learning scenarios. There are also a number of simulation-based approaches, e.g., *Making History* [16], or *SimCity* [17]. While simulation systems represent a promising tool to further learning, these games typically do not integrate easily into educational processes and also demand a careful application.

In the next chapter we will discuss the extent to which such commercial games meet the demands of learning and which learning processes have to be considered.

2 Integrating Learning into Computer Games

Schank and Cleary [18] consider learning processes to be successful if they address our natural way of learning. In their opinion such natural learning processes consist of three steps: Adopting a goal, generating questions, and developing an answer. They further propose several approaches to support such natural learning processes. These approaches are:

- (Simulation-based) learning by doing, where a student is actively engaged in a task and learns related skills in the course of performing his tasks,
- Incidental learning, where students acquire otherwise dull information while fulfilling a fun and interesting task,
- Learning by reflection, where students are stimulated to muse about a situation or a strategy applied to a task,
- Case-based learning which supplements the learning by doing approach by providing adequate expert knowledge in different situations when performing a task,
- Learning by exploring, where students' questions are answered and information is provided in a conversational format.

While Schank and Cleary provide examples on how to support such teaching approaches using information technology, the implementation of constructivistic approaches to learning on a broad level can still be considered an unsolved problem.

Requirements that result from the combination of game play and learning
GBL has been around for quite a while now, and several products are available that take on this approach. However, most of today's commercial learning games can hardly be considered convincing examples of successful game-based learning solutions. In general, these applications fall into two classes: those stressing learning, and those putting the focus on the game idea. While applications from the first class usually lack the targeted levels of fun and playfulness, applications from the second class typically fail to integrate learning and learning elements successfully. In general, the question on how to integrate learning successfully with elements of play and games is unsolved.

Jantke stated several requirements for educational games that most of today's GBL application do not fulfill [11]. Among these the following two demands were not met in most cases:

- Frustration caused by a conflict between game play and teaching material should be avoided e.g. learning interactions may not interrupt the flow of the game play and disturb the player's immersion. These interactions of learning have to be supportive instead and shall not hinder the player from reaching his goals.
- In general, the playing part and the learning part of the game should form a unit. The most important criterion is that interactions of learning should appear as inherent constituents of the game.

In the following we propose Interactive Digital Storytelling (IDS) to be utilized for GBL. As well, we discuss how IDS techniques can be used to interweave elements of game and learning and the requirements for such a system. We also discuss current examples and how far learning can be integrated into IDS applications.

3 Utilizing IDS for Game-Based Learning

A new approach is to combine verbal conversation and narrative storytelling principles in interactive computer environments. Stories not only represent the oldest cultural technique to convey information, they can also be understood as a central element of human thinking and communication [19]. Naturally, storytelling has been proposed as a principle for designing information systems [20]. Conversational elements supporting a natural dialogue between a user and the system represent a natural extension to this paradigm. Moreover, stories and conversational elements directly connect to Shank and Cleary's principles for successful learning scenarios.

Storytelling and conversational user-interfaces may also represent a starting point for rethinking e-learning concepts. Language is the medium in which we exchange thoughts and compare notes. In the classroom, language is most important. For example, the teacher phrases problems, explains algorithms and expresses ideas.

Unfortunatly, language software environments allowing interactive dialogues, as well as sophisticated computer games, easily get very complex, and extensive knowledge from various disciplines is needed for conceiving and building interactive storytelling systems. The current state of the art in interactive storytelling has a high focus on developing automated systems. It is driven by computer science research, especially in the fields of corpus-based natural language and AI planning algorithms.

Principles of Interactive Digital Storytelling

The concept of IDS has the potential to become a paradigm for future interactive knowledge media. It couples dramatic narrative with user interaction, providing highest forms of engagement and immersion. It also stands for the connection of games and stories by utilizing inherent structural elements of both.

Artificial characters taking the role of actors within a plot play an important role in the concept of Interactive Storytelling. Considering this, IDS agents can achieve more than simply being single virtual guides and virtual tutors, which are commonplace today in a variety of software products. As in stories, their role could be to interact with each other as a set of characters to present a dramatic storyline; and as in games, they have the potential to serve as all sorts of sparring partners for players to interact with, such as representing the bad guys, or companions who ask for help.

Detached from the actual content, there are design problems to solve concerning the dynamics of real-time IDS systems. At the same time, drawing from dramatic storytelling principles provides a set of experiences on the design of conversations in a way to fit the target group and to provide entertainment and fun. Spierling and Iurgel described open design issues that a storyteller is confronted with when designing and modeling an interactive story that involves several characters [21]. Here, the limitations of today's IDS systems represent a major challenge, and it is often very difficult to implement design concepts. A platform for interactive dialogues has to support characters with personality, story models to decide on possible plots and their connection to learning concepts, and the definition of interactions that are integrated within the written dialogue. Unfortunately, systems providing all of this functionality do not exist yet.

Successful implementations of intelligent conversations with animated virtual characters are rare, and there are no real success stories for such applications on the entertainment market to date. One of the few examples examining a middle course between the two approaches of linear stories and emergent behavior still is M. Mateas' and A. Stern's Façade [22]. It is based on a specialized dialogue management system and allows users to participate in a predefined and pre-recorded conversation between virtual characters. However, the system's design is focused on a specific scenario and authoring is currently supported for programmers only.

art-E-fact [21] and Scenejo [23] present similar integrations of simulation and plot. In contrast to Façade, an authoring system is central to the way a story is built. In art-E-fact, storywriters define digital conversations starting with a story graph of explicit dialogue acts, similar to branching, and provide more complex interactions by adding rules and chatbot patterns within nodes of the graph. With *Scenejo*, we take a different approach. Here, we start with chatbot text patterns to provide free conversational interaction with users. These patterns are merged into a story graph in a second step, allowing writers to line up conversational scenes and their parameters. Comparing *art-E-fact* with *Scenejo*, the first one is taking a top-down approach focusing on the story arc, while the second one takes a bottom-up approach emphasizing dialogue and interaction.

Scenejo is our technological approach to develop a framework for interactive storytelling applications for learning. *Scenejo* enables such playful simulations of dialogues between several conversational agents and multiple users. As mentioned above, *Scenejo* employs animated virtual characters and current chatbot technology as

the basis for text-based interactions. The result is an emerging dialogue, influenced by the users' inputs and the bots' databases of possible lines matching a text pattern coming from either a user or another bot. The bots also take into account parametric settings and scene descriptions provided by an author.

The experience of interacting with the platform of multiple chatbots shows that there is high entertainment value through the fact that the course of the conversation cannot be completely anticipated, not even by the writer of the dialogue patterns. While there are still problems with non-sequitur verbal reactions to user input, people mostly cope with it as within chats in their real life, and as a result, rather assume strange character traits to the bots according to their appearance.

Interactive Storytelling and Game-based Learning

At the moment there are only few games for learning implemented on IDS frameworks. *FearNot!* [24] is an interactive drama drawing from concepts of role-playing and improv theatre. It has been applied to education against bullying for primary-age children. *FearNot!* has a generic architecture (*ION*) and empathic character architecture (*FAtiMA*) that could be reapplied to other domains. Children playing this game are confronted with different bullying scenarios. They are not the victims themselves, but a virtual character. This character asks the learner for help, that means the child has to decide on how to react. Thus, the child learns by reflecting upon the situation and by identifying with the victim.

Another example is the *Killer Phrase Game* [25] based on the *Scenejo* platform. It tackles the topic of how to identify and react to so-called "killer phrases" within a discussion. The designed game assumes a scenario with two parties of planners and residents, arguing about novel plans for an airport extension. The partly predefined conversation between the two parties contains killer phrases. The learner plays the role of the moderator and has to manage the meeting (see Fig. 1).

This approach is a first step towards the utilization of IDS techniques for game-based learning. Thereby, predefined dialogs (elements of a superior story plot) between the learner and the virtual characters are used to convey information to the learner. This means, story and dialog are applied to achieve a learning goal, in this example to identify killer phrases.

At this point of the development, many technical aspects need improvement. But besides that, several strengths and opportunities, as well as weaknesses and risks, were identified [25]. During the design phase, an obvious learning effect has been the increased reflection upon the underlying model by the game designers and creators of the dialogues. The experiment has shown that there is a potential for designing successful games for learning on the basis of IDS.

Nonetheless, further development in the frameworks and the building of showcases is necessary, and will have to be evaluated. The question remains why IDS is a promising approach and why stories are good for learning.

Learning from Stories

Explanations for the capability of stories to enhance learning processes are quite complex and enfold the fields of narratology, (depth) psychology, and constructivism. Therefore, an all-embracing discussion would be too extensive in the context of this paper and the following elaborations do only reflect an extract.

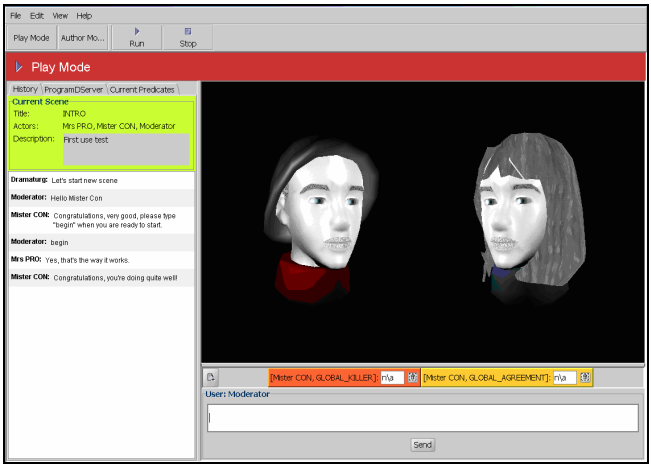


Fig. 1. Screenshot showing the main application window of *Scenejo* currently demonstrating the *Killer Phrase Game*

In every culture word-of-mouth stories function as a knowledge transfer between individuals, groups or generations. This transferred knowledge is not limited to facts, but also includes implicit cultural values, opinions, emotions, and solutions. Also, a narrative consists of a sequence of events, a timeline and a linear language representation. Successful stories are based on a traditional dramatic structure of contents which is rooted in old myths (e.g. [26] or [27]) and, allegeable regarding depth psychology, are associable with emotional necessities. Spierling wrote that learning elements, thus, can be conceived emotionally so that they are easier and more intuitively understandable than the complete complexity of system knowledge with several dependent variables [28].

Stories are a perfect element to communicate in contextualized knowledge. Hereby, provides the mechanism necessary to integrate into games.

Constructivistic Case-based Learning

The *Killer Phrase Game* is a first promising example for interactive storytelling applications in the field of learning. It fulfills requirements from the field of problem-based learning, as well as it shows the potential to implement instructional methods from cognitive apprenticeship.

Cognitive apprenticeship represents an approach in the context of the constructivistic paradigm. It represents a synthesis of formal schooling and traditional apprenticeship. The cognitive apprenticeship approach extends problem-based learning and situated cognition especially in terms of the processes of scaffolding and fading. Exactly these processes are very difficult to implement without the use of a personalized intelligent tutor. While such an intelligent tutor in the context of a cognitive apprenticeship approach comes close to one of an intelligent tutoring system, additional requirements exist. Scaffolding and fading in the sense of cognitive apprenticeship require a much more elaborate planning of temporal workflows, and, eventually, also

a definition of a dramatic arc. This provides a link to the field of narrative and to innovative approaches in the field of IDS.

Stories aren't Games and Games aren't Stories
At this point we would like to clarify the intersections and differences between computer games and interactive storytelling. Games provide the concept of play in a structured and goal-oriented way. In contemporary game theory there is an ongoing debate between ludologists (e.g., [29]) and narratologists (e.g., Brenda Laurel or Janet Murray), in which ludologists seem to aim at examining the game-specific dynamics of games [30], while Laurel, Murray, and others explore similarities and continuities by examining games through existing media like theatre or film etc. In the current discussion the position comes up that this debate might be a non-issue [31]. Maybe it is possible to agree on a few positions.

- Games do not necessarily involve story.
- Core differences at the centre of these two phenomena exist.
- Story can add to the pleasure of game-play.

Of course, games and story can certainly be combined, potentially resulting in a stronger user experience. But, they are not the same thing. Bizzocchi [32] believes that two areas have a lack of conceptual clarity: the concept of immersion and the term "narrative arc". Careful design of this arc is a powerful tool for channeling and guiding the experience of story. But the difficulty is that this depends on tight control over the design. And this is what highly interactive games do not afford. Nevertheless, IDS approaches exist that try to combine a narrative structure and an interactive process.

4 Requirements for Effective Learning Applications Based on the Interactive Digital Storytelling Paradigm

As mentioned above, the experiences with IDS applications are limited; those in the field of GBL even more. In fact, we are facing a chicken-and-egg problem here. On the one hand, it is difficult to design and realize IDS applications since suitable design strategies and development tools are missing. On the other hand, we do not really know which strategies are appropriate and how such development tools should look like since we are facing a new genre – if not a brand new media – with very new and unknown prospects [33].

Nevertheless, it is possible to point out some problem fields connected to some possible research directions. These problems and challenges can be classified as follows:

- Paradigms for interactive storytelling,
- Integration of games and play with IDS and learning,
- Interaction, and
- Technologies and standards.

In the following, we will discuss these aspects in some more detail.

Paradigms for Interactive Storytelling

IDS relies on the concept of a large number of interwoven story paths as opposed to a single story line in traditional media. While such multi-branch stories could be represented in terms of a story graph with multiple branching points determining the story, this approach is obviously not practical: the number of necessary branching nodes to achieve a convincing level of interaction and control by the user is just too large to be modeled and managed manually. Currently, two different approaches can be distinguished to solve this problem. Emergent narrative (e.g., [33], [24]) targets to provide users with a maximum of "choices" to steer the story into all possible directions. In corresponding systems the story is typically generated based on some kind of simulation based on actors with specific characteristics and goals, and their interactions. On a technical level, agent-based systems are often being applied in this context. While this approach is haunting, there are problems. It is very difficult to force some kind of dramatic structure in such an environment due to the lack of control on a general level. As such, experiences in such environments often resemble those in improvisation area, where interesting, unexpected events and boring sequences without dramatic content are equally probable.

On the other hand, guided interactive narrative (e.g., [34], [23]) prioritizes a coherent and interesting plot, following dramatic principles and drawing from experiences in other media such theatre and film. One approach in this field is based on the identification of story patterns in a sense similar to design patterns [35]. On a theoretical level, some of this work takes on classifications of stories and their elements such as Propp [27] and Barthes [36]. Corresponding story engines try to generate interesting stories from such story elements, following narrative principles. However, mapping narrative theories from linear media and the integration of user interaction proved to be difficult, and corresponding applications lack the level interaction one wood would expect.

The solution may lie in an integration of both concepts. Spierling coined the term "implicit creation" for this [37]. On a technical level, however, it is still not completely clear how such integration could look like.

Both strategies – emergent and guided interactive narrative – represent conflicting top-down approaches to the development of IDS systems. Probably it might also be worth rethinking the appropriateness of these top-down approaches in general, taking a different point of view. Schank [19] pointed out much earlier the function of stories as central elements in human thinking and communication. From his point of view, dialogue partners in human communication are always merely looking for a story to tell back. In the context of IDS, such stories could be understood as parametrizable micro-stories, consisting of a dramatic structure on a much lower level. Following this idea, storytelling could be understood as an intelligent selection of appropriate micro stories in a dialogue. This would correspond to a bottom-up approach. In terms of flexibility and choices such an approach could achieve similar flexibility compared to emergent narrative approaches. Moreover, control over the overall story and dramatic arc might still be possible to obtain. Advantages could especially arise within the context of GBL, where the integration of micro-stories might be easier to achieve than general concepts of guided interactive narrative.

Integration of Games and Play with IDS and Learning

In general, learning requires a much tighter control, which makes approaches from guided interactive narrative more attractive. Still, there are examples that apply solutions from emergent narrative in this field. Convincing GBL applications based on stories have to provide a good balance between story and game. However, this balance is not easy to achieve. Clearly, there are two approaches: a) Development starts with a game idea and design, where story and learning is integrated in a second step, or b) one starts with story and learning, and game-elements are added to this later on. Both approaches have their problems. Examples from the first class tend to have problems to integrate learning in overall game play, leading to interrupts in the game flow, and, consequently, to dissatisfaction for the player. On the other hand, experiences from the development of GBL applications show that the products tend to be boring and failing to provide a coherent game character if the design of central game elements is postponed [38]. Today, it is still unclear how to achieve the necessary integration.

This aspect is also closely connected to the aspect of authoring, and the development of appropriate authoring tools might be a first step to provide a solution.

Interaction

Interaction represents a further challenge in this field. Dialogue-based interaction seems the most natural approach. A number of systems allow for dialogues between users and virtual actors based on utterances (e.g., [22], [23]). Often, it is abstained from applying speech recognition, reducing interaction to text input and output, sometimes enhanced by speech output and elements of non-verbal communication based on animation. With the advancement of corresponding technologies, multimodal interaction and tangible device are being applied more often (e.g., [39]). Besides the general problem of speech and utterance recognition also the generation of utterances for virtual actors appropriate for a narrative context and reflecting the characteristics of a virtual character (knowledge, style, emotional state, etc.) remains a major challenge [40]. Both, enhanced input and output technologies are not only necessary to provide a coherent immersive experience in the game, but also to support dialogues on the necessary level of quality in the context of learning, where user input (e.g., questions, answers) have to be recognized in a reliable way.

Technologies and Standards

While a large number of game platforms became available in the last years, simplifying the development of computer games, for both, IDS and GBL corresponding platforms do not yet exist. Most IDS engines have still an experimental character. Often, they are targeted for the support of a single specific application, only. However, there are examples to integrate IDS into game engines (e.g., [41]).

Even less clear is how learning technologies such as learning management systems (LMS) could be linked to game-based learning systems. On a first glance, such integration might look strange, since the approaches to learning are very different, with courseware focusing mostly on factual and conceptual knowledge, while GBL targets to convey procedural and also implicit knowledge.

There are, however, good reasons to pursue such a connection. On the one hand, it can make sense to integrate GBL components as learning objects into course materials

managed by an LMS. This would require SCORM compliance and interoperability of the GBL components. This includes the acquisition of student performance data from a GBL component for assessment purposes. On the other hand, the more or less well-structured learning content available in LMS can be seen as a valuable source for background material and explanations in GBL scenarios.

Standardization will be necessary also on other levels to ease the development of GBL applications by supporting the exchange of plots, characters (geometric models, animations, personalities, etc.), dialogues, and simulations.

5 Conclusions

The integration of Interactive Digital Storytelling with game-based learning has the capability for creating effective educational computer games. We have to stress that Interactive Digital Storytelling is about the development of a new genre. At present, we do not know what good interactive stories are. Also, the experiences with IDS in the field of Game-based learning are very limited. First experiments are however promising. The *Killer Phrase Game*, a small and simple educational game on rhetoric and communication could be implemented successfully based on the Scenejo framework. So far, it has shown that there is a potential for designing successful games for learning involving virtual actors based on digitally implemented agents. But there is a need for further technical development and a lot more of convincing showcases have to be built. However, these open tasks are preconditions for applying IDS successfully in the field of learning.

References

1. Prensky, M.: Digital Game-Based Learning. McGraw-Hill, New York (2001)
2. Gee, J.P.: What Video Games Have To Teach Us about Learning and Literacy. Palgrave Macmillan, New York (2003)
3. Lepper, M.R., Chabay, R.W.: Intrinsic Motivation and instruction: conflicting Views on the Role of Motivational Processes in Computer-Based Education. Educational Psychologist 20(4), 217–230 (1985)
4. Rieber, L.P.: Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. Educational Technology Research & Development 44(2), 43–58 (1996)
5. Gee, J.P.: Situated language and learning: A critique of traditional schooling. Routledge, London (2004)
6. Malone, T.W., Lepper, M.R.: Making learning fun: A taxonomy of intrinsic motivations for learning. In: Snow, R.E., Farr, M.J. (eds.) Aptitude, Learning and Instruction III: Conative and Affective Process Analyses. Erlbaum, Hillsdale (1987)
7. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. Harper Perennial, London (1990)
8. Bowman, R.F.: A “Pac-Man” theory of motivation: Tactical implications for classroom instruction. Educational Technology 22(9), 14–16 (1982)

9. Dickey, M.D.: Game design narrative for learning: Appropriating adventure game design narrative devices and techniques for the design of interactive learning environments. Educational Technology Research and Development 54(3), 245–263 (2006)
10. Mishra, P., Foster, A.N.: The Claims of Games: A Comprehensive Review and Directions for Future Research. In: Crawford, C., et al. (eds.) Proc. of Soc. for Information Technology and Teacher Education International Conf., pp. 2227–2232 (2007)
11. Jantke, K.P.: Games that do not exist communication design beyond the current limits. In: Proc. ACM Conference on Design of communication (2006)
12. Jenkins, H., Squire, K.: Harnessing the power of games in education Insight 1(3), 5–33 (2004)
13. Van Eck, R.: Digital Game-Based Learning: It's Not Just the Digital Natives Who Are Restless. EDUCAUSE Review 41(2), 16–30 (2006)
14. Westermann: Mathematikus (2008) (last visited: 28.02.08), http://www.westermann.de
15. Klett, H.: Informaticus (2008) (last visited 28.02.08), http://www.braingame.de/
16. Muzzy Lane Software, Making History (2007) (last visited: 28.02.2008), http://www.making-history.com
17. Wright, W.: SimCity (1989) (last visited 28.02.08), http://www.maxis.com/
18. Schank, R.C., Cleary, C.: Engines for Education. Lawrence Erlbaum Pub., Mahwah (1995)
19. Schank, R.C.: Tell Me A Story – Narrative and Intelligence. Northwestern Univ. Press, Evanston, Illinois (1995)
20. Murray, J.H.: Hamlet on the Holodeck: The Future of Narrative in Cyberspace. The Free Press (1997) ISBN 0-684-82723-9
21. Spierling, U., Iurgel, I.: Just Talking About Art - Creating Virtual Storytelling Experiences in Mixed Reality. In: Balet, O., Subsol, G., Torguet, P. (eds.) ICVS 2003. LNCS, vol. 2897, pp. 179–188. Springer, Heidelberg (2003)
22. Mateas, M., Stern, A.: Integrating Plot, Character and Natural Language Processing in the Interactive Drama Façade. In: Proc. TIDSE 2003, Darmstadt, pp. 139–151 (2003)
23. Weiß, S.A., Müller, W., Spierling, U., Steimle, F.: Scenejo – An Interactive Storytelling Platform. In: Proc. ICVS 2005, Strasbourg, France, pp. 77–80 (2005)
24. Aylett, R.S., Louchart, S., Dias, J., Paiva, A., Vala, M.: Fearnot! - an experiment in emergent narrative. In: Panayiotopoulos, T., Gratch, J., Aylett, R.S., Ballin, D., Olivier, P., Rist, T. (eds.) IVA 2005. LNCS (LNAI), vol. 3661, pp. 305–316. Springer, Heidelberg (2005)
25. Spierling, U.: "Killer Phrases": Design steps for a digital game with conversational role playing agents. In: Mayer, I., Mastik, H. (eds.) Proc. Isaga 2007, Eburon Pub. (2008)
26. Campbell, J.: The Hero with a Thousand Faces. Princeton University Press, Princeton (1949)
27. Propp, V.: Morphology of the Folktale, 2nd edn. University of Texas Press (1968)
28. Spierling, U.: Interactive Digital Storytelling als eine Methode der Wissensvermittlung. In: Eibl, et al. (eds.) Knowledge Media Design. Oldenbourg Verlag (2006)
29. Frasca, G.: Ludology Meets Narratology: similitudes and differences between (video) games and narrative. Originally published in Finnish as Ludologia kohtaa narratologian in, Parnasso, 3 (1999), English version online at http://www.ludology.org
30. Frasca, G.: Simulation 101: Simulation versus Representation (2001) (last visited: 28.02.08), http://www.ludology.org/articles/sim1/simulation101.html
31. Pearce, C.: Theory wars: An argument against arguments in the so-called ludology/narratology debate. In: de Castell, S., Jenson, J. (eds.) Changing views: Worlds in play, Digital Games Research Association Conference Proceedings, Vancouver, BC (2005)

32. Bizzocchi, J.: Games and Narrative: An Analytical Framework. In: Proc. CGSA 2006 (2006)
33. Crawford, C.: Chris Crawford on Interactive Storytelling; New Riders. Berkeley (2005)
34. Szilas, N.: Interactive Drama on Computer: Beyond Linear Narrative. In: AAAI Fall Narrative Intelligence Symposium. AAAI 1999 Fall Symposium Series (1999)
35. Alexander, C.: A Pattern Language. Towns, Buildings, Construction. Oxford University Press, New York (1977)
36. Barthes, R.: Introduction to the structural analysis of narratives. In: Sontag, S. (ed.) A Barthes Reader. Hill & Wang (1966)
37. Spierling, U.: Adding Aspects of "Implicit Creation" to the Authoring Process in Interactive Storytelling. In: Cavazza, M., Donikian, S. (eds.) Virtual Storytelling, Proc. ICVS 2007. LNCS. Springer, Heidelberg (2007) (Best Paper Award)
38. Naone, E.: Grandios gescheitert. Technology Review (January 2008) (last visited: 28.02.08), `http://www.heise.de/tr/artikel/100944`
39. Cavazza, M., Lugrin, J., Pizzi, D., Charles, F.: Madame Bovary on the Holodeck: Immersive Interactive Storytelling. In: Proc. ACM MULTIMEDIA 2007, pp. 651–660 (2007)
40. Cavazza, M., Charles, F.: Dialogue Generation in Character-based Interactive Storytelling. In: AIIDE 2005, pp. 21–26 (2005)
41. Charles, F., Mead, S.J., Cavazza, M.: Generating Dynamic Storylines through Characters. Int. Journal of Intelligent Games & Simulation (IJIGS) 1 (March 2002)

Designing Virtual Players for Game Simulations in a Pedagogical Environment: A Case Study

Jean-Marc Labat

Lip6 Lab, UPMC (Paris 6)
104, Av. du Président Kennedy, 75016 Paris, France

Abstract. The development of learner' activities is a key element in the design of a pedagogy based on constructivism. A classic way to implement this pedagogy in practice consists in using simulation. When the simulation is a game, the learner is stimulated by competition with other learners. But, sometimes, there are not enough human players. In order to increase the "playability", we need to introduce virtual players. These virtual players must be defined with respect to 4 properties: (i) to play in a normal way, neither too well nor too poorly, (ii) their behaviour must be unpredictable, (iii) they must not cheat and (iv) they must not be distinguishable from human players. In this paper, we propose a methodology to define such virtual players and we illustrate it in the case of the SIMPLUS project, a business game.

Keywords: Learning Environments, virtual player, business simulation, reverse-engineering, knowledge-based system.

1 Introduction

As de Jong proposed [Calie 04], learners are encouraged to construct their own knowledge in realistic situations. This supposed an increasing of the learner's activities. There are several ways to improve learner's activities. On the one hand, the networks enhance communication among learners and also with the tutor whose role has been acknowledged as highly important. On the other hand, it is important that, whenever possible, the learner be engaged in problem-solving activity involving realistic situations. This indeed is the case when the tutoring system includes a problem solver, an environment based on microworld, a virtual reality environment or a simulator [Guéraud et al., 99].

In this paper, we focused on particular simulators which are business games. These simulation activities are widely used in France at college level (particularly in engineering and business schools) to train students in company management.

In section 1, we analyse briefly the different kinds of simulation and the level at which students interact. Section 2 provides a general description of a particular business game and of the architecture of the system. Section 3 presents the pedagogical environment that the company, with which this project is developed, needs. Their need was to introduce virtual players in order to increase the gameplay when there are not enough human players. As we are in an commercial context, the aim of our industrial partner is to provide a large set of business games on line in an ASP context.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 487–496, 2008.

Therefore, our goal is not to define virtual players for one particular game but rather for a host of games. Section 4 presents a more or less generic methodology for designing virtual players. We underline the epiphytic character of the pedagogical environment developed, using the metaphor proposed by [Giroux et al., 96]. Lastly, we present some results of experimentations and related works.

2 Simulation and Pedagogical Environment

As it is said by [Beaufils and Richoux, 03], "Simulation software can be considered as suited for learning theories...but there is a wide range of simulation software and information on how to use it in science teaching/learning is often missing". They distinguish three types of activities related to the manipulation of models: (i) the activity of modeling itself, which is often used in physics or chemistry, (ii) the manipulation of models where the learner simply enters data and looks at the result and (iii) the discovery of a complex model which is the most characteristic and interesting activity. The goal is that students can infer non elementary properties which are consequences of the model but which are not in the model itself. It is generally the case when the simulation software is a game because, if the results are too simple, such a game is not entertaining enough. For example, in our case, which is a business game, the link between the values the students give to the decision variables like the price of their products, and their shares of the market, is not at all obvious.

But, this is not sufficient to be sure that the students will improve their knowledge and their understanding of the management of a company. Pedagogical functionalities such as explanations or diagnosis of what they have done must be added to the simulation software to create the conditions of learning. Moreover, the pedagogical environment must provide help to the tutor [Gouarderes et al, 99]. In our case, the pedagogical environment must provide the possibility for the tutor who pilots the simulation to introduce virtual players in order to enhance the learning of the human players. Here, we do not present the explanations part but only the designing of virtual players.

3 An Online Business Simulation Game

Working in the context of a project named SIMPLUS and in collaboration with an industrial partner EXOSIM, we focused on business games.

3.1 Visual-Surf, a Business-Simulation Game

Visual-Surf is an online simulation game developed by EXOSIM. It is a business game involving a wintersports-material company.

Each team represents a company producing snowboards, funboards and surfboards. We use the term team because at the college level, students are gathered to constitute teams of three or four persons. But, today, they discuss the decisions they are about to make out of line. The environment does not provide any functionality to collaborative learning. Therefore, in this paper, team and player are synonymous. As it is the game, the tutor who supervises the simulation is sometimes named the referee.

The game is composed of ten periods, each corresponding to a business year for the company. In each new period, the playing team must make a general decision about all the different parameters of the company before a deadline which is determined at the beginning of the simulation. Generally, the tutor allows fifteen days for the decision process.

This decision is composed of production decisions (management of the production equipment), commercial decisions (management of the commercial and marketing parameters of the company), and financial decisions (budget and portfolio management). During the game, it is possible for the teams to communicate with each other and also with the tutor. In this way, the referee can advise the teams, and the teams can exchange information relevant to subcontracting and tendering (take-over bids).

At the end of each period, the simulation engine run with all the decisions received from the teams. Then the results are sent to all the teams together with the referee's comments. The results obtained give a broad outline of the company and its competitors. These results belong to different categories: general results common to all the companies (market share per product and company; average investment by company in advertising, research, quality, organisation and maintenance; average price by product for all companies…); accounting results which are specific to each company (income statement; assessment; details of stock by product in volume and value); production results specific to each company (management of the production equipment: assignment of production units, detailed cost price; calculation of the production cost of the products).

After the last period, a winner is designated on the basis of his score. Apart from the first decision, there are three sources of information for each team: the historical logfile of the past decisions, present results of the team and its competitors, and a market study provided by the system. With this information, the player can define his own strategy and plan a new decision.

3.2 General Description of a Multi-player Simulation Game on the Web

EXOSIM has developed an online business-simulation game. Two websites have been created: one for the players and the other for arbitration. The first site presents

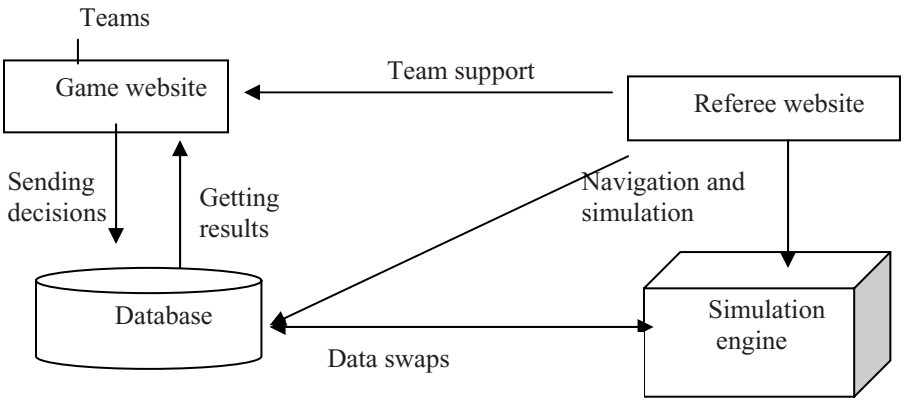


Fig. 1. The system architecture

all the services of the game: participation in the contest, presentation of the results, comments, e-mails and FAQ. The second site is accessible only to the referees.

Initially, the referee plays the role of the organiser: he defines the simulation according to different scenarios, fixes the complementary parameters and manages the simulation (data processing, virtual players …). Secondly, he plays the role of tutor: he analyses the decisions and the results of the teams, gives them advice and also answers the e-mailed questions. As [Crampes and Saussac, 99] wrote, "the choice of the scenario and the technical and human organisation are fundamental". Concerning the technology developed in HTML and ASP, the sites run on a SQL Server database and a simulation engine developed using Visual Basic.

4 A Methodology to Design Virtual Players

4.1 Needed Characteristics of Virtual Players

If the virtual player's behaviour is too different from that of the human player's, the simulation will not be attractive enough; this will not only detract from the playability of the game but will also, probably, divert the students from the subjacent objective which is to acquire concepts of business management. It is thus essential that the virtual players be accepted by the human players. Bearing this in mind, it is necessary that they have the following characteristics:

1. They must play in a normal way, neither too well nor too poorly. This, however, is not necessarily easy to arrange. If the games are too simple, for example if they are algorithmic in nature, the virtual players can play perfectly. On the other hand, if the games require real expertise, the virtual players may have aberrant behaviours, in particular when they are designed along symbolic lines.
2. Their behaviour must be unpredictable so that, in identical situations, the human players cannot anticipate their reactions. This is *a priori* easier to arrange (you just need to introduce more randomness into the game). However, in most games (console games) the behaviour of the virtual players is completely stereotyped which means that their human adversary can learn their reactions "by heart". In addition to making the game more random, it is possible to define several categories of virtual players, as we propose in this paper (see below).
3. They must not "cheat". On the one hand, if the human players notice that the virtual players are not playing fair, they will probably feel distaste for the game; moreover, this cheating can hinder the training. Indeed, if the virtual players are efficient because they possess information which they are not supposed to have, the learners may make false conjectures about the virtual player's reasoning[1]. Consequently, they must be no doubt about the information available to the virtual players. In this particular case, the virtual players have the same information as the human players, but the expertise model they contain is the same as the model which is implanted in the simulation. This is one of the fundamental ideas of this

[1] This can be seen in computerized bridge games which use knowledge of the four deals to play the cards. This leads to aberrant behavior (for instance, making a useless finesse because the players know it will work).

project: collecting the expertise of a field is not difficult insofar as, by definition, this expertise is in the simulation engine. To meet the different needs of a simulation, the virtual players were designed as an expert system (SE). The expert-system approach was the obvious choice as it can represent a true game rationality: i.e., adapt the decision to the economic situation in the game (company results, market trends etc.), evolution of the decision-making (change of strategy or behaviour during the game)… Thus the virtual player is not simply a calculus algorithm. He is a module composed of a set of objects and rules, a module which sets the numerical parameters of the virtual player's decisions.

4. An epiphytic[2] virtual player: In order to adapt to several simulations, the virtual player must be able to integrate easily into the online game. The expert-system approach respects this criteria: the CLIPS module is integrated into the data-processing environment of the decision through a DLL enabling the virtual player to communicate with the database and also with the website used by the referee (with this website, the referee can define the number of players, their behaviour and some global economic variables about the game). The expert system operates in the same way as the human player with his calculus sheet and his decision sheet. It finds all the necessary information for decision-making (results of the previous period and parameters of the current period), it processes it and sends the numerical data of the new decision to the database. Its running does not interfere with the functioning of the game: the expert system must not modify the data processing of the simulation. *For human players* and *for the system*, there is no difference at all between the decisions taken by virtual or real players: The same procedure enters the data into the database. The only difference comes from the fact that the procedure receives data from a form for the human players and from the expert system for virtual players. Thus, the virtual player can be described as an epiphytic system [Giroux et al., 96].

4.2 A Methodology Based on Reverse Engineering

It is well known that to explicit the expertise is the bottleneck for the development of expert systems, especially when the goal is to model humans. But, in the case of simulation, we are in a very favourable situation as the expertise of the domain is represented within the simulation engine in a procedural manner. So, what we have to do when defining virtual players is only to represent this expertise in a declarative manner. Moreover, as human players are represented in the database by attributes, we use the same attributes, with the same names, to describe virtual players in the expert system, using the Clips templates. These two points are the key points when defining the project. As we will see in section 4.3, we define a methodology in four stages.

4.3 Stages

1 Stage 1: Analysing the game and modelling the decision-making

The first objective of the project is to create virtual players which can be used in several simulations. This stage makes it possible to answer the fundamental questions

[2] In botany, an epiphyte is a plant which grows on another plant without disturbing it in any way.

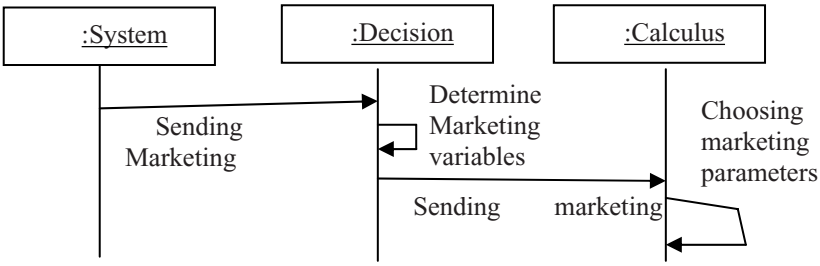


Fig. 2. An example of UML Models

which are raised by the creation of the virtual player: In which environment does the virtual player operate? What are the rules of the game? What are the strategies of the game? How should he play? The aim is to build a general outline of a player's decision-making based on two identification processes: identification of the various steps of the decision (which makes it possible to define the way to be followed by the virtual player to make his decision) and identification of the decision variables of each step, which makes it possible to define the numerical parameters. Modelling was done according to the UML method (see figure 2). The virtual player was constructed in two modules in order to define a reliable game behaviour: the first module represents the decision making of the player and the second the procedural aspect; i.e. how the simulation engine computes the parameters.

2 Stage 2: Selecting in the database the necessary tables and the attributes to represent the players and the expertise

The data generated by the decision-making of the human players are stored in a database and are treated later by the simulation engine. The analysis of these data makes it possible to determine the relevant tables and the right attributes representative of the player's decision and to transform them into an object structure within the expert system. The results of the simulation engine after each period are also stored in the database. The relevant attributes to model the expertise of the game also were transformed into object structures within the expert system. This point is essential for the reverse engineering methodology. From our experience, there exists a large similarity between the Entity/Association model for the designing of database and the modelling of objects in expert system. It is even simpler in expert system as there are, very often, multivalued attributes.

3 Stage 3: Analyzing the functions used by the simulation engine

From the analysis of the functions used by the simulation engine, it is possible to define two kinds of functions: calculation functions which will be reproduced in the expert system just as they are in the simulation engine, and the functions which model the expertise (identification of the parameters of the functions and extraction of the limit values used by the engine to define the different behaviours of the player).

For example, when the player determines his cost price, he should first of all set certain parameters such as publicity. Let X be, the value of the parameter publicity fixed by the player, and Y(x) the value of a marketing variable calculated by the engine from publicity, the simulation engine uses an interval of value [Ymin, Ymax] to

which must belong Y(x): if Y(x)>Ymax then Y(x)=Ymax. When these parameters are set, the cost price is automatically calculated by calling upon calculation functions which require no decision.

4 Stage 4: Designing the virtual player

Once the information necessary for the decision-making has been determined, it is then necessary to develop the expert system representing the player. It must meet the needs defined in stage 1 and present different possible behaviours. One of the objectives in the development of the virtual player was to be able to choose between different game attitudes.

A cautious player: He is prudent in his approach towards market shares and never takes great risks while managing his company. Thus, he arranges things so that he will not borrow too much and always keeps enough factories. He invests just the right amount in research. With regard to prices, he always sets high margins in order to be on the safe side.

A neutral player: He is typically average. He invests just the right amount in research; he produces neither too much, nor too little. He sets margins that are not particularly high but not very low either. Finally, the neutral player does not have any ambition.

An ambitious player: He is the complete opposite of the cautious player. He aims to get the highest market share; he accepts large loans and small margins in order to crush competition. He invests huge amounts to be sure to produce surfboards as soon as possible.

These three types have been described by coefficients which are internal to the expert system. These coefficients are based on calculations relative to the decision and on the threshold values of the simulation engine. These values determine intervals to which the results of the decision made by the player belong: the aggressive player is close to the upper limits and the fearful player to the lower limits.

5 Some Results

An interface was developed to accomplish preliminary tests. The performances of the expert system were compared to the decisions made by the players whose strategies were similar to the three behaviours defined above. The tests were carried out on three types of periods, each of which has a specificity in the time-course of the game: the tests of the period 1 type were chosen as they involve the initialisation of the simulation, the tests of the period 2 type because the companies cannot yet produce all the products (they are no surfboards), and the tests of the period 7 type because the companies which have completed their research phase can produce surfboards. Initially, each player (cautious, neutral and ambitious) was tested over these periods in the CLIPS environment. After that, the test interface was used to test them in the simulation environment. The results showed that the expert system took into account the various behaviours in a random way: a cautious player never made the same decision in two consecutive episodes, and this was true for the two other types of players as well.

The first contest with the virtual players was carried out with eight players: five human players and three virtual players. We implement the three types of virtual player: cautious, ambitious and neutral. Among the human players, there were the two designers of the simulation game whose objectives were to test and evaluate the virtual players (and the generated explanations). The last three players were students who compete in normal conditions. Therefore, the two designers made voluntary mistakes to see how the system reacts.

Another important point to notice is the fact that one contest is not necessarily representative because there is a great variability in the game: first of all, the tutor chooses the scenario (for example, he decides of the evolution of the market), secondly there is a random part in the choices made by the virtual players and in the simulation itself, thirdly, the tutor chooses the criteria used to rank the players.

Keeping in mind these preliminary remarks, the results were very satisfactory. The ranks of the virtual players were two, four and five. Moreover, the comments of the designers of the game, who participated in the contest, were positive. They said that virtual players had an average level: their decisions were good but they did not anticipate as humans would. It is exactly what we want, as the virtual players must be neither too good neither too bad. Since this first experimentation, current uses of the simulation game confirm these first results: the virtual players are reasonably good and the students cannot distinguish them from real players.

6 Related Work and Discussion

Since more than fifteen years, many works introduced virtual agents in Intelligent Tutoring System. Generally speaking, there are two different possible goals: either the objective is to simulate a co-learner, i.e. a companion who learns with the human learner; either the objective is to simulate pedagogical agents, that is to say, virtual agents who play the roles of a teacher or a tutor. This is the case in classical intelligent tutoring systems, where a pedagogical agent could serve as an expert tutor to teach knowledge to learners [Koedinger & Anderson, 97]. An example is the Adele agent [Johnson et al., 00], who represents an expert in the domain of medicine. On the other side, others researchers have suggested that agents could serve in instructional roles such as learning companion [Chan & Chou, 97], collaborator [Dillenbourg & Self, 92], competitor [Chan & Baskin, 90], or even trouble maker [Aimeur & Frasson, 96)]. Learning Companion Systems (LCSs) extend the traditional model of ITSs by adding computerised agents whose aim is to provide a peer for the human student. This kind of agent is called a *LearningCompanion* [Chan, 91]. It acts as a teachable student of the human student. [Uresti, et al. 05] have developed a LCS to explore the hypothesis that a learning companion with less expertise than the human student would be beneficial if the student taught it. The system implemented two companions with different expertise and two types of motivational conditions. They observe that students in the motivated condition with a weak companion taught it many more times than in the other experimental conditions and in general worked harder.

Another approach is proposed by [Baylor and Kim, 04] who distinguish three particularly roles for a pedagogical agent. They design Expert, Motivator, and Mentor roles for college students within the MIMIC (Multiple Intelligent Mentors Instructing Collaboratively) agent-based research environment.

- The function of the expert agent was to provide accurate information in a succinct way. He looks like a traditional teacher.
- a motivator agent who spoke enthusiastically and energetically. He was presented as an eager participant who suggested his own ideas, verbally encouraged the learner to sustain at the tasks, and, by asking questions, stimulated the learners to reflect on their thinking. He expressed emotion that commonly occurs in learning, such as frustration, confusion, and enjoyment [Kort et al., 01]. The motivator is a kind of learning companion.
- A mentor should be a guide or coach with advanced experience and knowledge that can work collaboratively with the learners to achieve goals. Thus, the mentor should simultaneously demonstrate competence to the learner and develop a social relationship to motivate the learner [Baylor, 00].

In this work, the goal of the virtual players is not at all to be a learning companion or another kind of pedagogical agent. The goal of the virtual players is only to maintain the gameplay, as the industrial partner realized, through some contests, that when the number of players is not enough, the motivation of the teams decreases and, consequently, they learn less than when the competition is more attractive. In our case, the learning is widely implicit. During the play, only the explanations given after each episode are explicit learning. The virtual players only play an indirect pedagogical role that is to increase the gameplay and thus the motivation. They do not interact with the other players, neither as a companion, neither as a mentor. Therefore, the approach taken in this paper is quite different from those usually taken in tutoring system. The goal of the design of virtual players is reached. Nevertheless, it could be interesting to design an expert agent and/or a mentor agent to give the explanations at the end of each episode in a way more attractive than it is for the moment.

7 Conclusion

The development of virtual players and explanations based on reverse-engineering technology using the expertise implanted in the simulation, seemed to be effective: the virtual players are correct, even though they can be slightly improved. But, as the company died in 2006, we have not tested the methodology with other simulation games in order to measure the generic part of our methodology.

Finally, we are sure that simulation-based learning will be more and more developed in the future as it allows to increase the learner‘s activities, using the dynamic capacities of computers. We totally agree with [de Jong T., 04], constructivism is nowadays one of the dominating learning theories, at least in Computer Based Learning context.

References

[Aimeur, E. & Frasson, C. 96] Analyzing a new learning strategy according to different knowledge levels. Computers & Education 27(2), 115–127 (1996)

[Baylor A. & Kim Y. 05] Simulating Instructional Roles through Pedagogical Agents. International Journal of Artificial Intelligence in Education 15, 95–115 (2005)

[Baylor, A. L. 00] Beyond butlers: Intelligent agents as mentors. Journal of Educational Computing, Research 22(4), 373–382 (2000)

[Beaufils & Richoux, 03] Un schéma théorique pour situer les activités avec des logiciels de simulation dans l'enseignement de la physique, Didaskalia. pp 9–38 (2003)

[Chan, T.-W. 91] Integration-Kid: A Learning Companion System. In: Mylopolous, J., Reiter, R. (eds.) Proceedings of the Twelfth International Conference on Artificial Intelligence (IJCAI-1991), vol. 2, pp. 1094–1099. Morgan Kaufmann Publishers Inc., San Francisco (1991)

[Chan, T. W. & Baskin, A. B. 90] Learning companion systems. In: Frasson, C., Gauthier, G. (eds.) Intelligent tutoring systems at the crossroads of artificial intelligence and education, pp. 7–33. Ablex Publishing Corporation, NJ (1990)

[Chan, T. W. & Chou, C. Y. 97] Exploring the design of computer supports for reciprocal tutoring systems. International Journal of Artificial Intelligence in Education 8, 1–29 (1997)

[Crampes M. & Saussac G, 99] Facteurs qualité et composantes de scenario pour la conception de simulateurs pédagogiques à vocation comportementale. revue STE 6(1), 11–36 (1999)

[de Jong T., 04] Learning Complex Domains and Complex tasks, the Promise of Simulation Based Trainig. In: CALIE 2004 conference, Grenoble, pp. 17–23 (2004)

[Dillenbourg, P. & Self, J. 92] People power: A human-computer collaborative learning system. In: Frasson, C., McCalla, G.I., Gauthier, G. (eds.) ITS 1992. LNCS, vol. 608, pp. 651–660. Springer, Heidelberg (1992)

[Dupuy J.-P, 99] Aux origines des sciences cognitives, édition La Découverte, Paris, p. 188 (1999)

[Giroux S., Paquette G., Pachet F., & Girard J., 96] EpiTalk - A Platform for Epiphyte Advisor Systems Dedicated to Both Individual and Collaborative Learning Intelligent Tutoring Systems, 363–371 (1996)

[Gouarderes, G., Minko A., & Richard L., 99] Simulation et environnement multi-agents pour l'apprentissage de la maintenance d'avions. revue STE 6(1), 143–187 (1999)

[Guéraud V., Pernin J-P., Cagnat J-M., & Cortes G. 99] Environnements d'apprentissage bases sur la simulation- Outils auteur et experimentations. revue STE 6(1), 95–141 (1999)

[Johnson, W. L., Rickel, J. W., & Lester, J. C. 00] Animated pedagogical agents: Face-to-face interaction in interactive learning environments. International Journal of Artificial Intelligence in Education 11, 47–78 (2000)

[Koedinger, K.R. & Anderson, J.R. 97] Intelligent tutoring goes to school in the big city. International Journal of Artificial Intelligence in Education 8, 30–43 (1997)

[Kort, B., Reilly, R., & Picard, R. W. 01] An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion. In: Proceedings IEEE International Conference on Advanced Learning Technologies (2001)

[Uresty J. & du Boulay B., 04] Expertise, Motivation and Teaching in Learning Companion Systems. International Journal of Artificial Intelligence in Education 14, 67–106 (2004)

The Relationship between Game Genres, Learning Techniques and Learning Styles in Educational Computer Games

Kowit Rapeepisarn[1], Kok Wai Wong[1], Chun Che Fung[1], and Myint Swe Khine[2]

[1] School of Information Technology, Murdoch University,
South St, Murdoch Western Australia 6150
{k.rapeepisarn,k.wong,l.fung}@murdoch.edu.au
[2] Emirates College for Advanced Education, United Arab Emirates
mskhine@westnet.com.au

Abstract. Educational computer game has many similar characteristics like any other genres of games. However its particular aspect is designed to teach, and in which main objective involves in learning a topic. To develop an effective educational computer game, different game genres, learning activities and techniques, and learning styles are important issues for consideration. This paper presents an analysis by comparing and establishing relationships between the game genres and learning techniques based on the types of learning and potential game style of Prensky [1] and learning styles based on the study of Chong et al. [2].

1 Introduction

Educational computer games and various forms of edutainment have gained much attention in the discipline of learning and teaching. Educators [3], [4], [5] believe that most children learn best through play. Most studies also show that 'learnt through play' [6] has proven to be a successful learning experience. Therefore, it is desirable to use educational computer games for teaching, which carry the objectives of play and learn in the classrooms. There are reasons for using computer games as a learning tool to enhance the learning experience of students. These reasons include the incorporation of rules, goals, engagement, challenge, feedback, fun, interactive, outcome and immediate reward [1], [7], [8], [9]. Even though most genres of the computer games in some ways are educational, educational computer games are designed with explicit educational purpose. When educational computer games are adopted in supporting learning in the classroom, the pedagogical aspects such as learning style should be taken into account. As different people learn and process information differently, it is important to understand individual learning style which allows the prediction of the way learners react and feel in different situations. Selecting the appropriate game genres for learning is another important issue for consideration to develop effective educational computer game. Recently, most studies focus on several variables when selecting game genres. This includes age level, gender, racial diversity, number of players, and the role of teacher.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 497–508, 2008.

Unfortunately, previous studies on the learning styles of the learners have not provided sufficient guidelines to design effective educational computer games to meet the needs of individual learners. However, there are two important studies which are relevant to the focus of this paper: Prensky [1] and Chong et al. [2]. Prensky's study presented a theory based on computer games and learning, whereas Chong's study focuses on the impact of learning styles using digital games. However, there are still some gaps between their works. Therefore, this study aims to explore alternatives by focusing on the learning techniques and the learning activities to match possible game genres discussed in Prensky and Chong's experimental research on computer game types and the four learning styles. In addition to the literature review, this paper takes a further step to develop a conceptual model based on those two studies in order to make a contribution to the knowledge.

2 Why Use Game for Learning?

Majority of children today are growing up in a digital society. Being accustomed to digital technology, children have changed considerably their ways of thinking and processing information based on different mindsets from their parents. For most children, computer games have become a major part of their lives and become the most common activity in children's leisure time. To help in understanding the differences between today and previous generations, and to justify why computer games need to be a part of education, Prensky [1] proposes ten aspects of comparison that include: 1) twitch speed and conventional speed, 2) parallel processing and linear processing, 3) graphic first and text first, 4) random access and step-by-step, 5) connected and standalone, 6) active and passive, 7) play and work, 8) payoff and patience, 9) fantasy and reality, and, 10) technology-as-friend and technology-as-foe.

In the educational aspect, educators believe that children learn best when it is fun. It can be said that it is a natural way for children to learn through play. Through play, human can acquire skill without knowing it. Most studies [1], [3], [4], [5], [6] also show that "learn through play" is a natural and universal learning tool for children and adult. Therefore, it makes sense to see play as a valued contributor to a child's development and it should be given a place in the school curriculum. Computer game, a media that is based on playing and entertaining, can be treated as learning-oriented game which is also known as "edutainment" [5]. While edutainment bring the concept of entertainment and education at the same time, computer games also bring together the idea of game, play, fun, and hand-on experience in the learning environment. Consequently, playing computer game can be assumed that it is the activity of learn through play. Prensky [1] confirms two reasons why use computer games for learning: 1) today's learners have change radically, and 2) these learners need to be motivated in new ways. Furthermore, the main reasons people play games because the process of game playing is engagement and games bring combination of motivating elements [1].

Apart from these, there are several other reasons why computer games can be used as a learning tool. The reasons include: computer games have rules, goals, interaction, and content and story. Gee [4] mentioned that games are heavily motivating. They teach people to think about complex systems to solve problems in a complex world. Games make player think about decision they are making and how do the decision will

impact on this situations. Games deploy rich visual that draw players into fantasy worlds and motivate the player via fun, via challenge and via instant feedback. With instant feedback and immediate reward computer games provide, it makes a crucial aspect for learning.

3 Educational Computer Game

Basically, educational computer games have the same characteristics like any other types of computer games. The particular aspect of this type of game is designed to teach, and in which the main objective involves the learning of a lesson. Rather than being structured as a straightforward set of lessons or exercises, this type of educational software is structured like games, with such elements as scoring, timed performances, or incentives given for correct answers. Some examples of educational computer games include: Basic Math, eduProfix, Mario's Early Year, Fun with Numbers, Mario Teachers Typing, Math Blaster, Episode 1, Math Grand Prix, Morse, Number Games, Pelmanism, Playschool Math, Spelling Games, Urban Jungle, Word Games, Zoombinis.

Many educational games of the past have been skill-and-drill (the common example is MathBlasters) One could argue that there is a place for skill-and-drill in learning; other might suggest that educational games need to be built on constructivist or social constructivist theoretical frameworks [10]. When educational computer game is adopted as a learning tool in classrooms, teacher should either create or adapt the learning materials to maximize the game's potential to support learning. As such, the pedagogical value should definitely be taken into account when considering in adopting educational computer games for teaching and learning in a classroom. For example, computer game might be integrated into instructional design and should affect children's capabilities to perform certain cognitive functions [11]. From the study of Chuang and others [11], they found that cause-and-effect games tended to encourage means-end analysis strategy, whereas adventure games encouraged inferential and proactive thinking. Moreover, outcomes from several researches proved that significant correlation exists between game playing and children's problem solving skill and cognitive style [4], [12]. In order to make educational computer game "educational", Fisch [13] suggests that when designing game, the following matters should be included: 1) matching the education topic to the media, 2) placing educational content at the heart of the game play and 3) building feedback that supports learners into the handling of difficult content.

4 Learning Styles

Research into the use of games in education is growing rapidly [5], [7], [14], [15]. In order to understand the potential roles of mainstream games in supporting learning, we need first answer the questions "what is learning?" and "what forms of learning are suitable for incorporating games in the classroom?" This is related to pedagogical theory which includes learning theory which describing how people learn or what styles of learning people like. Learning style is useful in identifying the methods by which people prefer to receive information from their environment and undertake their

learning. Each person has his or her own way of converting, processing, storing, and retrieving information. Some people prefer to learn through reading and reflecting on how this might apply to their own experience, whilst others prefer to learn through trying ideas out and learn through reviewing their experience before planning the next step. Among the learning styles which classified as experiential, Honey and Mumford learning style is one of the well known experiential learning [16]. This learning style proves that people learn better when the teaching is adapted to the learning styles [2]. Honey and Mumford classify learners into activist, reflector, theorist, and pragmatist as illustrated in Table 1.

Table 1. Honey & Mumford learning style [16]

Characteristics			
Activists	**Reflectors**	**Theorists**	**Pragmatists**
• Immerse in new experience • Enjoy here and now • Open minded, enthusiastic, flexible • Seek to centre activity around themselves	• Stand back and observe. • Cautious, take a back seat • Collect and analyse data about experience and events, slow to react conclusion • Use information from experience to maintain a big picture perspective	• Think in a logical manner, rationally and objectively. • Assimilate facts into coherent theories. • Fit things into rational order. • Keen in basic assumptions, principles, theories, models and thinking system.	• Keen to put ideas, theories and techniques into practice. • Search new ideas and experimental • Act quickly and confidently on ideas, get straight to the point. • Are impatient with endless discussion

5 Game Genres, Learning Techniques and Learning Styles

Game genres can be categorised as Action, Adventure, Fighting, Puzzle, Role-Playing, Simulation, Sports, Strategy, etc. Different game genres have different impact on the content of learning activities [1]. Some contents are best learned through role-playing and adventure games, others are best through game show competition, action and even sport games [2]. Different games appeal to different people. Choosing the appropriate type depends on the content to be learned and /or mental processed to be developed. Prensky [1] proposes several variables to consider when selecting a game style including: target age level, gender, racial diversity, number of play, and the role of the teacher. Apart from these, pedagogy aspect especially learning style should be considered as one important variable. Knowing children's learning style and finding appropriate ways to create or enhance learning environment such as choosing appropriate game type for each style of learners will increase the student's learning success. Consequently, the questions follow are "what types of game should be created for each learning technique?" and "what types of game are appropriate for each learning style?"

This section presents the model of relationship between learning activities and game types based on the study of Prensky [1] in section 5.1; learning style and education game based on the study of Chong et al. [2] in section 5.2; This paper then proposes a

new conceptual model by comparing and matching learning styles, learning activities and game genres based on those two studies [1], [2] in section 5.3.

5.1 Prensky's Study: Learning Techniques and Learning Activities Used in Educational Computer Games

Games can be categorised in many different genres, including first-person shooters, role-playing, action, adventure, card, puzzle, and sports. If computer games are used in the classroom, the game genres should be selected to match the learning style. Different activities and learning techniques may use different types of game. Prensky

[1] shows activities and learning techniques used in educational computer games are: 1) practice and feedback, 2) learning by doing, 3) learning from mistake, 4) goal-oriented learning, 5) discovery learning and guided discovery, 6) task-based learning, 7) question-led learning, 8) situated learning, 9) role playing, 10) constructivist learning, 11) multi-sensory learning, 12) leaning objects, 13) coaching, and, 14) intelligent tutors.

In his paper "Computer Games and Learning: Digital Game-Based Learning" [1], Prensky discusses about how to combine gameplay and learning. He claims that teacher have to understand the types of learning content. With different kinds of learning content, teacher can see what kinds of learning are really going on such as learning fact, skill, judgment, theory, reasoning, process, procedure, creativity, language, system, observation and communication. Additionally, teacher can choose different learning activities according to particular types of content. Prensky proposes the relationship of learning content, learning activities and possible game style as shown in Figure 1 and Table 2.

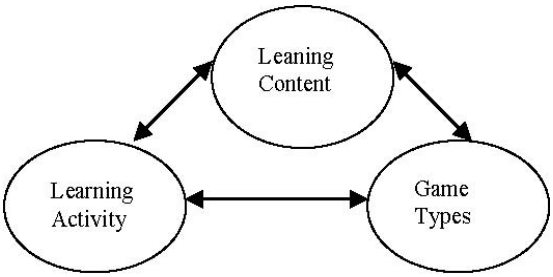


Fig. 1. Model of relationship of learning content, learning activities and possible game styles

5.2 The Study of Chong et al.: Learning Styles and Educational Game

Many educational researchers studied on learning styles, but the study on the relationship of learning styles to learning within game text is scarcely found. Researchers are emphasizing that the education computer games should be developed considering the learning styles of students [2]. However, there is a research conducted by Chong et al. on the impact of learning styles on the effectiveness of digital games in education. They conducted the survey based on the Honey and Mumford four types of learning styles on 50 undergraduate students in INTI College Malaysia. They choose three

Table 2. Finding Summary of Prensky's Learning Content, Leaning Activities and possible Game Styles [1]

Learning Content	Learning Activities	Possible Game Styles
Facts : laws, policies, product	Questions, memorization, drill, association	Game show competitions, flashcard types game, mnemonics
Skills: interviewing, teaching, management	Imitation, feedback, coaching, continuous practice	Persistent state games, role-play game, detective games
Judgment: management, decisions, timing, ethics	Reviewing cases, asking questions, feedback, coaching	Role-play games, multiplayer interaction, ad-venture game, strategy game, detective game
Behaviors: supervision, self- control, setting example	Imitation, feedback, coaching, practice	Role-play game
Theories: marketing rationales, how people learn	Logic, experimentation, questioning	Open ended simulation games, building game, construction games
Reasoning: strategic & tactical thinking, quality analysis	Problems, examples	Puzzles
Process: Auditing, strategy creation	System analysis & deconstruction, practice	Strategy games, adventure games
Procedure: assembly, bank teller, legal	Imitation, practice, play	Timed games, reflex games
Creativity: invention, product design	play	Puzzles, invention games
Language: acronyms, foreign language	Imitation, continuous practice, immersion	Role-play games, reflex games, flashcard games
Systems: health care, markets, refineries	Understanding principles, graduated tasks	Simulation games
Observation: moods, morale, inefficiencies, problems	Observing, feedback	Concentration games, adventure games
Communication: appropriate language, involvement	Imitation, practice	Role-play games, reflex games

different kinds of games namely: Counter Strikes, Championship Manager and Bookworm which are action role-playing game, strategy game and puzzle game respectively. The results show the student's preferences on the games vary related to learning styles. Chong et al. concluded that they need to conduct further studies on different types of learning styles as well as different game genres. The finding of Chong's study can be summarised in Table 3.

5.3 Bridging the Gap between the Prensky and Chong's Studies

When reviewing the studies of Prensky and Chong at al., we realized that more need to be done in order to provide a better framework for designing good educational games. Prensky focuses on learning techniques, learning contents, and learning activities but lack of learning style. Whereas, Chong et al. focus on learning style but uses only three

Table 3. Experimental finding summary based on Chong et al. [2]

	Role-playing games (Counter Strike)	Strategy games (Championship Manager)	Puzzles (Bookworm)
Activists	Enjoy playing this game	Discard the instructions given before the start of the game	Use their brainstorming to solve the problem
Reflectors	Prefer not to lead the game	Observed to follow the instructions given to them earlier	Not able to draw strong conclusion
Theories	Not able to draw strong conclusion	Reacted very similar to the reflectors	Did not learn and play well
Pragmatists	Dislike this game	Copied the strategy given during the briefing	Great interest in this game

different game genres as an example. Therefore, it is the attempt of this study to bridge this gap by establishing the linkage of these two studies. Two conceptual models proposed in this study are:

Firstly, as mentioned in section 5.1, Prensky proposes the relationship of learning content, learning activities and the game styles. He also suggested 14 essential learning techniques which he claimed that it should be considered and used when designing learning materials. However, these 14 learning techniques have not been matched to learning activities and game genres in his study. In order to use all those learning techniques in learning with educational computer game, the relationship of each learning techniques and game genres should be studied. Hence, it is the objective of this paper to compare and match his 14 learning techniques to learning activities and game genres. The new model and the result of this matching are illustrated in Figure 2 and Table 4.

Second, in section 5.2 Chong et al. [2] study the impact of learning styles on computer game in education. They use three types of games in their study as the examples to prove that different learning styles do prefer different types of games. From the finding of their study, the behaviors of each style of learners while playing game are also described. However, they do not match this behavior of each style of learners with learning technique. Additionally, only three different game styles are studies. While the study of Prensky choose all the standard categories of computer games matching with learning activities, but lacks of the comparison of learning style of the users. Thus, this paper proposes the new conceptual model of the relationship between learning styles, learning activities, and possible game genres based on these two studies [1], [2] as illustrated in Figure 3.

The process that led to the new model is conducted by: 1) exploring the behavior when playing games for each type of learner based on Chong's study, 2) matching the behavior of each type of learner to learning activities based on Prensky's study, and 3) finding the possible game genres which can relate to each learning activities. As an example, the results from this study found that the possible game genres for the activist learners could be multiplayer interaction, action games and role-playing game. Accordingly, the results of matching learning styles; behavior when playing game; behavior when using computer; learning activities; and possible game genres are shown in Table5.

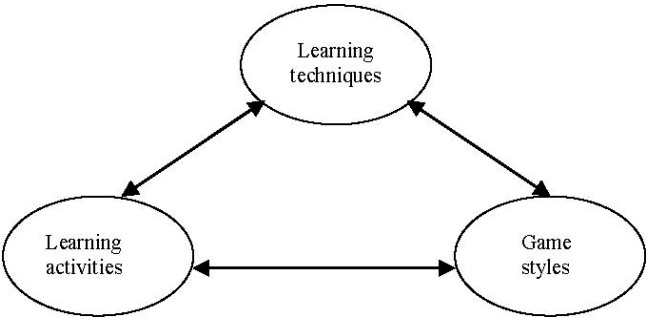


Fig. 2. Model of relationship of learning techniques, learning activities and possible game styles

Table 4. The relationship between learning techniques, learning activities and possible game styles

Learning techniques	**Leaning activities**	**Possible game genres**
Practice & feedback	Questions, memorization, association, drill, imitation	Game show competition, flashcard type game, mnemonics, action, sports game
Learning by doing	Interact, practice, drill, imitation	Strategy game, action game, role playing game
Learning from mistake	Feedback, problem	Role-play game, puzzle game
Discovery learning & guided discovery	Feedback, problem, creativity play	Adventure game, puzzle game
Task-based learning	Understand principle, graduated tasks	Simulation game, puzzle game
Question-led learning	Question/ questioning, problem	Quiz or trivia game, game show competition, construction game
Situated learning	Immersion	Immersive style game such as role-playing game, flashcard game
Role playing	Imitation, practice, coaching	Role-playing game, strategy game , reflex game, adventure game
Constructivist learning	Experimentation, questioning	Building game, constructing game
Multisensory learning	Imitation, continuous practice, immersion	Game in which introduce new technologies such as locatable sound or force feedback, reflex game
Learning object	Logic, questioning	Games which are becoming object-oriented
Coaching	Coaching, feedback, questioning	Strategy game, adventure game, reality testing game
Intelligent tutors	Feedback, problem, continuous practice	Strategy game, adventure game, puzzle game, reflex game

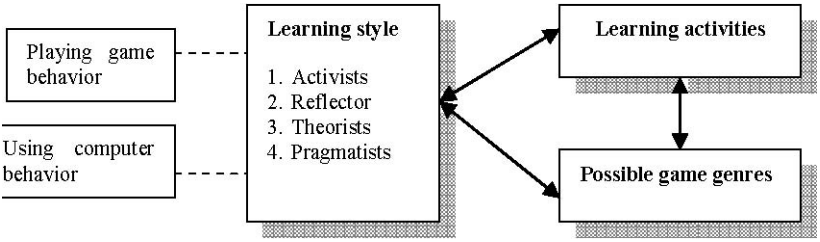


Fig. 3. Model of relationship of learning styles, learning activities and possible game genres

6 Discussion

As researchers have found that computer games have significant educational value, computer games can become part of the school curriculum. There are different types of computer games and games technologies, which have been used positively, both directly and indirectly to support and assist teaching and learning in the classroom. Green and McNeese [17] conclude in their paper "Using Edutainment Software to Enhance Online Learning" that the attributes of high quality educational computer games comprise of 1) clear learning goal and objective, 2) provide review on concepts newly learned and allow for questions and answers, 3) develop higher thinking skills, 4) challenging but focus on learning rather than on winning or losing, 5) clear rules so learners know how to play, 6) providing a means for collaboration, feedback, or guidance, 7) be fun so learners are more relax, more alert, less fearful and open to learning, 8) provide a means for debriefing to recap what was learned and allow for question and answer. However, there are a number of issues, which need to be addressed before using computer games in the classroom.

Most studies concern about age of student, gender, racial diversity, and role of the teacher. Unfortunately, not many researches focus on learning style when designing the appropriate game genres for each style of learners. Different people have different style of learning. No single learning preference is better than any other. In fact, individual student may have more than one single learning style. This reflection can be shown that when some learners prefer kinesthetic instruction, they can also have ability to learn orally and visual [18]. One learner has the active type of learning; he/she may have theory or pragmatic style of learning in other learning situation. Therefore, there are many possible ways in choosing appropriate game genres for one particular student. However, it can be determined by looking at learner's dominate learning style. In other word, the most preferred learning style of that learner. The model of relationship of learning style, learning activities and possible game genre presented in this paper is only the potential example proposal. To understand educational gaming, many factors have to be examined. These include design, pedagogy, and literacy. It should also focus on the classroom use, what is learned and what can be taught with educational computer game [13]. Moreover, some variables such as the experience in playing game, culture, language, and nurture should also have been examined. Game developers and educational psychologists should work together with other professionals as a team to formulate the educational content in order to build quality educational computer games.

Table 5. The relationship of learning style, behaviour when playing game, behaviour when using computer, learning activity and possible game genres [1], [2], [16]

Leaning styles	Behaviour when playing game	Behaviour when using computer	Learning activities	Possible game genres
Activists	Prefer working as a team, being a group leader, Be able to brainstorm to solve the problem	Like to use shortcut key-combinations but will also find the toolbar buttons useful.	• Practice • Imitation • Work with other • Tackle problem 'head on'	Multiplayer interaction, action game, role-playing game
Reflectors	Go through the important data in	Prefer to use dropdown menus but	• Observing • Feedback	Concentration game,
	the game, follow the instructions, spend a long time before make decision, not to lead the game	will soon discover what is best for themselves, like to browse through SEARCH FOR HELP in the HELP menu	• Graduated task • Work alone at their own pace	adventure game, simulation game
Theorists	Go through the data and follow the instruction before start the game, be able to give careful thoughts when choosing the game elements, Formulate good strategy to defeat the enemy	Often use dropdown menus to see what else the application can do, like to browse through the INDEX or SEARCH FOR HELP in the HELP menu	• Logic • Understanding Principle • Analyse & develop plan • Explore relationship between things	Strategies game, simulation game
Pragmatists	Follow closely the instructions & strategies that were mentioned in the briefing, believe they can play better if they were given proper instruction, Show a great interest in puzzle game and dislike role-playing game	Probably use the toolbars buttons to get things done, often find HELP menu to get things done	• Experimentation • Asking question, • Try things out • structure plan with definable purpose	Puzzle game, building game, constructing game, reality testing game, detective game

The further questions are: How do educators convince parents, teachers, and administration about the importance of gaming in education? [13]; Are there any significant gains when determine the appropriate genre of game with learning styles? And are there any relationship between teaching styles and using of educational computer games?

7 Conclusions

Educational computer games bring together the idea of game, play, fun, hands-on experience and also with explicit educational purpose. Like other genres of computer games, educational computer games have elements that benefit learning. This includes rules, goals, active engagement, content/story, feedback, interactive, problem solving, quick adaptation and immediate reward. However, educational content in education computer games should be considered as the heart of game, and feedback that support learners should also be built into difficult content. When designing educational computer games for supporting learners in classroom, apart from educational content, they can also embrace the pedagogical benefits such as learning style with game genres for developing quality learning experience in class. This paper shows the comparison and matching of learning techniques, learning activities, learning styles to possible game genres. However, this study merely proposes the potential model used as one possible conceptual guideline for further study in order to create effective educational computer games. Further study may find out the benefit gain when determine the appropriate genres of game with learning styles.

References

[1] Prensky, M.: Computer Games and Learning: Digital Game-Based Learning. In: Raessens, J., Goldstein, J. (eds.) Handbook of Computer Game Studies, pp. 97–122. The MIT Press, Cambridge (2005)

[2] Chong, Y., Wong, M., Thomson Fredrik, E.: The Impact of Learning Styles on the Effectiveness of Digital Games in Education. In: Proceedings of the Symposium on Information Technology in Education, KDU College, Pataling Java, Malaysia (2005)

[3] Morgan, A., Kennewell, S.: The Impact of Prior Technological Experiences on Children's Ability to Use Play as a Medium for Developing Capability with New ICT Tools. ICT Research Bursaries, Becta (2005)

[4] Gee, J.: What Video Games Have to Teach Us About Learning and Literacy. Longman, New York (2003)

[5] Kirriemuir, J., McFarlane, A.: Literature Review in Games and Learning `http://www.futurelab.org.uk/download/pdf8/research/litreviews/Games_Review1.pdf` (accessed in November 2005)

[6] Rapeepisarn, K., Wong, K., Fung, C., Depickere, A.: Similarities and Differences Between "Learn Through Play" and "Edutainment". In: Proceeding of the 3rd Australian Conference on Interactive Entertainment. Perth, Australia., December 4-6, pp. 28–32 (2006)

[7] Blunt, R.: A Causal-Comparative Exploration of the Relationship Between Game-Based Learning and Academic Achievement: Teaching Management with Video Games, Ph.D. dissertation, Applied Management and Decision Sciences; Walden University (2006)

[8] Garcia, G.: Digital Game Learning. In: Hoffman, B. (ed.) Encyclopedia of Educational Technology (2005), http://coe.sdsu.edu/eet/articles/digitalgamlearn/start.htm (accessed in November 2006)

[9] Juul, J.: The Game, the Player, the World: Looking for a Heart of Gameness. Level Up. In: Digital Game Research Conference Proceedings (2003)

[10] Ferdig, R.: Preface: Learning and teaching with electronic games. Journal of Educational Multimedia and Hypermedia 16(3), 217–223 (2007)

[11] Chuang, T., et al.: The Effect of Computer-Based Video Games on Young Children: A Literature Review. In: Crawford, C., et al. (eds.) Proceedings of Society for Information Technology and Teacher Educational International Conference, pp. 3954–3957. AACE, Chesapeake,VA (2005)
[12] Dreyfous, R.: Cognitive and Affective Variables Involved in Recreational Computer-Generated Games. Dissertation Abstracts International, 55(07), 1807A (UMI No. 9432617) (1994)
[13] Fisch, S.: Making Educational Computer Games. Educational, 56–61
[14] Ju, E., Wagner, C.: Personal Computer Adventure Games: Their Structure, Principles and Applicability for Training. Data Base for Advances in Information Systems 28, 78–92 (1997)
[15] Amory, A., et al.: The Use of Computer Games as an Educational Tool: Identified of Appropriate game types and game elements. British Journal of Educational Technology 30, 311–322 (1999)
[16] Honey, P., Mumford, A.: The Manual of Learning Styles. Peter Honey, Maidenhead (1992)
[17] Green, M., McNeese, M.: Using Edutainment Software to Enhance Online Learning. International Journal on ELearning 6, 5–16 (2007)
[18] Dunn, R.: Individualizing Instruction for the Mainstream Gifted Children. In: Milgram, R. (ed.) Teaching Gifted & Talented Learners in Regular Classrooms, Charles C. Thomas, Springfield, Ill (1989)

EFM: A Model for Educational Game Design

Minzhu Song and Sujing Zhang

College of Education, Zhejiang Normal University, 321004 Jinhua, Zhejiang, China
song-mz@163.com, sjzhang@china.com

Abstract. The research and development of educational game in our country is still in a primary stage, and effective models and ideas for educational game design are of great lack. The educational game can not only be considered as a kind of instructional media, but also as a games-learning environment. The paper proposed the EFM model for educational game design through describing the internal connection of motivation, flow, effective learning environment and educational game. Toward creating an effective learning environment, the EFM model aims at inspiring motivation through flow. Based on this model, some ideas are suggested, intending to provide some design guidelines for researchers and developers of educational game.

Keywords: Motivation, Flow, Effective Learning Environment, Educational Game.

1 Introduction

With the reform and development of education, contemporary educators increasingly concerned about the overall development of learners, who will be promoted as a separate entity for treatment. Under the guidance of such an educational idea, a vast number of educators and parents commonly concern how to inspire learners' motivation and help them truly learn from playing. In the modern society, with the popularity of network games and electronic games, the youth become infatuated with the computer games, which promotes some scholars more actively research how to turn the games effect into education. However, when many educational experts are concerned about the educational value of games, the games was just viewed as teaching media. In fact, the game can not only be regarded as a kind of teaching media, but also as a learning environment to study, because the game itself contains the basic elements which are necessary in learning activities. Through referring to relevant information at home and abroad, it is clearly found the close connection of motivation, flow, effective learning environment and educational game. Educational game can serve as an effective learning environment, so that learners produce motivation during the process of flow experience and change from passive learning to active learning, so as to enhance the quality of learning.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 509–517, 2008.

2 Related Theory

2.1 Motivation

According to the Great Chinese Dictionary (CiHai), motivation refers to the driving force in promoting people to learn. Whether the students study positively, what do they glad to study, and how do they study, all have direct relations with the motivation. As some scholars said, motivation can not only result in learning activities, but also enhance the efficiency and improve the effect of learning. The first step of instructional design is to consider how to inspire motivation and maintain it. To help understand motivation in instruction, we can look at the ARCS Model of Motivational Design as developed by John M. Keller of Florida State University. The ARCS Model identifies four essential strategy components for motivating instruction [1]:

[A]ttention strategies for arousing and sustaining curiosity and interest.
-Learners are more motivated when the instructional design generates curiosity and interest about the content or learning context.

[R]elevance strategies that link to learners' needs, interests, and motives.
-Learners are more motivated when goals are clearly defined and align with learners' interests.

[C]onfidence strategies that help students develop a positive expectation for successful achievement.
-Learners are more motivated when challenge is balanced in such a way that the learning process is neither too easy as to bore the leaner, or too difficult such that success seems impossible.

[S]atisfaction strategies that provide extrinsic and intrinsic reinforcement for effort.
-Learners are more motivated when there are rewards for correctly executed actions.

2.2 Flow

Flow Theory developed by Csikszentrnihalyi of Chicago University, as a method for understanding and implementing motivation, is a theoretical bridge between the concerns of instructional design and motivational design theory [2]. Flow theory has been widely applied in the discussion about behavior and psychology in the man-machine interaction environment, and it has been confirmed that the flow experience does exist in the use of network.

Csikszentrnihalyi (1991) defined the phenomena of flow state as having nine dimensions [3]:

- Goals of an activity.
- Unambiguous feedback.
- Challenge-skill balance.
- Action-awareness merging.
- Concentration.
- Control.
- Loss of self-consciousness.
- The transformation of time.
- Autotelic experience.

Based on the process of flow experience, Novak and others (2000) classified these nine dimensions into three categories [4]:

- Conditional factors, which could stimulate flow experience, include goals of an activity, unambiguous feedback and challenge-skill balance.
- Experience factors, the feeling of individuals in a state of flow experience, include action-awareness merging, concentration and control.
- Result factors, the results of experience of individuals in the flow state, include loss of self-consciousness, the transformation of time and autotelic experience.

2.3 Effective Learning Environment

The so-called learning environment is an integration of the supporting conditions which promote the development of learners. This shows the possibility and significance to create a learning environment [5]. Learning environment design is to create an effective and positive learning environment, in order to help the students understand and master the learning content, and improve the abilities of self-cognition [6]. Norman identified seven basic requirements of an effective learning environment [7]:

- Provide a high intensity of interaction and feedback.
- Have specific goals and established procedures.
- Motivate.
- Provide a continual feeling of challenge that is neither so difficult as to create a sense of hopelessness and frustration, nor so easy as to produce boredom.
- Provide a sense of direct engagement, producing the feeling of directly experiencing the environment, directly working on the task.
- Provide appropriate tools that fit the user and task so well that he can get aid and do not distract.
- Avoid distractions and disruptions that intervene and destroy the subjective experience.

2.4 Educational Game

Education and game are originally an indivisible whole. Especially for children, game is learning. Children's study start from observation, imitation and inquisition, and the best activity which can embody Children's spirit of learning is game. When the children play role playing games, play with sand or imitate adult activities they are interested, their active attitude, cooperation spirit, explore awareness, imagination, and even a certain degree of creative intelligence all will get trained and developed. "Learning from playing" has been stressing since the era of Confucius, and the principles of all learning activities modern educators advocated can be reflected in the game. The game is one of the important ways for children and adult to study.

"Educational game" is still a newly emerging thing in our country, and there is no explicit definition nowadays. Narrowly speaking, educational game refers to the integration of education and game, and the education effect naturally generated from the process of playing games, in other words, it means "a type of computer game software which generates education effect through interest" [8]. The educational game in narrow

can be defined as excellent educational game. Broadly speaking, educational game means all computer software which includes both educational material and game elements [8]. It includes electronic game clearly pointed to the education application, as well as some healthy electronic game with educational value or other study aids software with the effect of game.

This research points at the narrow definition of educational game, which is studied as a learning environment. It is considered that the educational game is a games-learning environment followed game mechanisms. Through the created situation and the internal rules, the educational game can stimulate participants' interest, as well as the expectation of final victory. At the same time, the game content is richer in knowledge and education effect. In the virtual challenging context, learners are required to learn and apply various skills and knowledge to complete the designed tasks, in order that they can acquire knowledge and skills, develop intelligence, cultivate emotions, attitudes and values, and achieve the purpose of education.

3 The Raising of EFM Model

Through the interpretation of the theories above, it is clearly found the connection of motivation, flow experience, effective learning environment and educational game, as shown in Figure 1:

Educational game can provide the virtual environment with specific targets and preset procedures to learners. Learners participate in the scenes, and challenge to specific tasks with existing knowledge, skills and appropriate tools. They can get access to the feedback during the interaction with the environment, adjust their behaviors and go on playing with the incentive mechanism. In this period, learners will almost not notice that they are going through a learning process in an urgent rhythm, and also can not describe the relevant principles and motives of their own action. In fact, they have already learned some knowledge or skills. Obviously, the educational game is an edutainment environment, containing many essential conditions of effective learning environment. Through rational design, it will have the completely possibility to become an effective learning environment.

As an effective learning environment, it must be provided with the seven basic requirements, which include providing students certain tasks with clear goals and appropriate challenges, and achieving a high degree of interaction and feedback. They are exactly related to the three conditional factors that stimulate flow experience. Obviously, in an effective learning environment, learners can certainly acquire a flow experience.

The four essential strategy components for stimulating motivation--relevance strategies, confidence strategies, satisfaction strategies, and attention strategies, relate to the four elements which are goal, challenge, feedback and interest. The nine dimensions of the flow experience also include clear goal, unambiguous feedback, challenge-skill balance, and concentration. Evidently, the state of the flow experience includes the four essential strategy components for stimulating motivation. Once a learner goes into the state of the flow experience, he will achieve positive study under the driving of the inherent motivation.

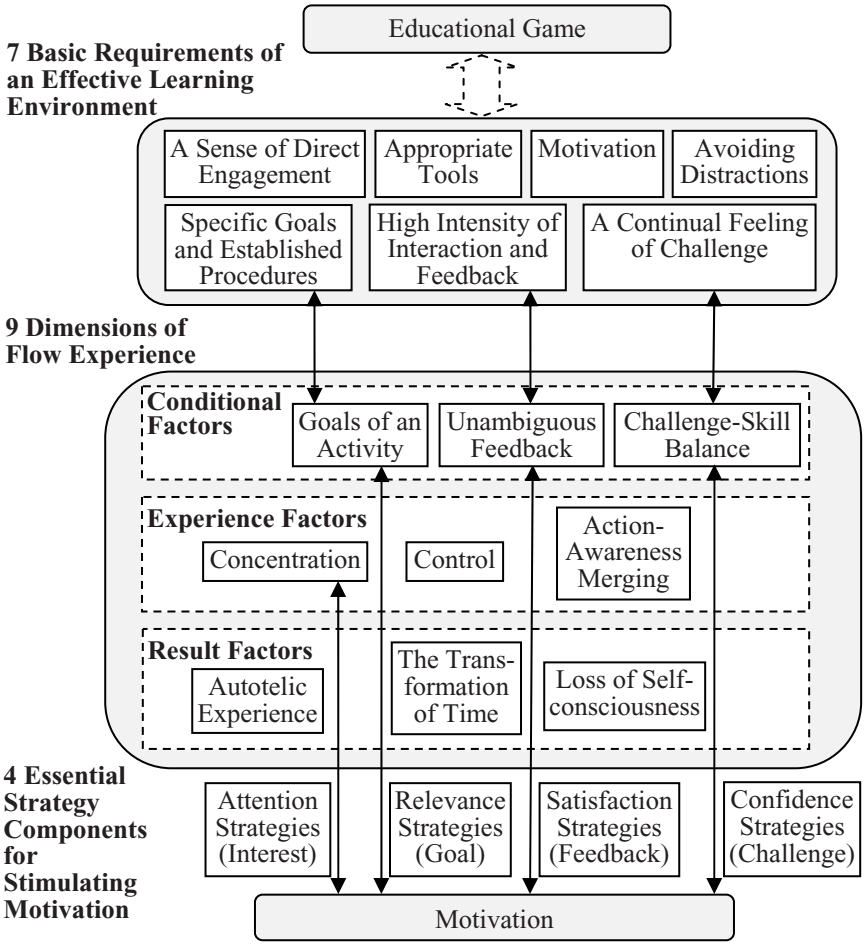


Fig. 1. The connection of motivation, flow, effective learning environment and educational game

In summary, learners can acquire flow experience in an effective learning environment, and the flow experience can certainly stimulate motivation. A well-designed educational game itself can be an effective learning environment for stimulating motivation and promoting learning.

Above all, "EFM: a model for educational game design" is raised. EFM is the acronym of effective learning environment, flow experience and motivation. Model is shown in Figure 2:

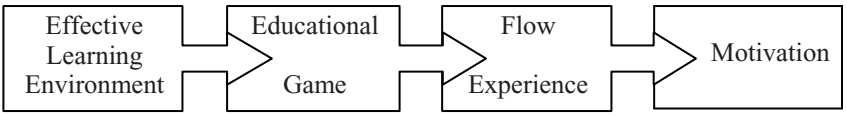


Fig. 2. EFM model for educational game design

According to the model, educational game can be treated as a learning environment. For the orientation of creating an effective learning environment, the educational game can be designed according to the prerequisites of an effective learning environment. In order to help learners acquire flow experience in the effective games-learning environment, inspire motivation, and improve the quality of learning.

4 Ideas of Educational Game Design on the Basis of EFM Model

Based on the EFM model, the designer can embark the educational game design from the fulfillment of the basic requirements of an effective learning environment. Therefore, the designer will, in the process of educational game design, explore how to construct the learning environment from the following aspects:

4.1 Specific Goals–To Set up the Goals of Educational Game According to 3D Objective

Educational game has educational aims or instructional objectives, which separate from or integrate with the original goals of game [8]. New curriculum standard emphasizes the 3D objective, as perfect combination of knowledge and skill, process and method, attitude and values, which is one of the important references to set up the goals of educational game. So it is very essential to emphasize the process of copious and natural experience of game, and the cultivation of correct passion, attitude and values in addition to the knowledge and skill. The goals of educational game should be set up in the principle of promoting full development of students.

4.2 Established Procedures–To Provide Learning Procedure through the Setup of Scenes and Rules

To ascertain the type and characteristics of game according to the corresponding curriculum and content first, and then confirm the frame of game scenes based on the detailed teaching unit. We can divide the content into several units on the basis of the goal of educational game, and then divide the game into several scenes on the basis of the content units. That is to say, when we finish the frame of game scenes and the relevant rules, we fulfill the whole game procedure and provide the learning procedure for the learner accordingly. In the game environment, the learner can do everything according to the scheduled procedure, and what he needs to do is to immersing himself into the game body and soul.

4.3 Appropriate Tools–To Provide Help Tools through Props

Suitable props should be provided to help the learner to fulfill the tasks so that he will not give up when he meet difficulties. Prop is a kind of important motivator in the game, which can awake the curiosity of the learner so as to enhance the recreation of the game. When the task comes with tools, the attention will be moved from learning goal to play, as a result, the learners can reach the state of flow experience.

4.4 Avoiding Distractions–To Avoid Distractions through Transparent Control of Game

All the available resource will be used to deal with the relevant information (main tasks) but not the game control because of the limitation of information treatment by human beings. Therefore, in an ideal situation, game control will be transparent, and the learner can be absorbed in the game.

It will destroy the first reception and interest to a game if there are long-winded instructions for the game operation and procedure. And if the learner needs to study how to play the game because of the complicated operation, it will occupy the attention and other cognitive resource of the learner, which will slow down the process of the game and even block the achievement of flow experience.

Consequently, the game interface should be concise and will not confuse the learner, the relationship and the layers of guide system should be clear and will not let the learners get lost, the definition of common function key should be similar or same to the common games, so that the learner can learn it easily. In conclusion, the transparency of game control can guarantee that the learner will pay maximum attention to the game and can reach the state of flow experience without any interference.

4.5 A Sense of Direct Engagement–To Enhance the Sense of Direct Engagement through the Real and Multi-choice Plot

The backgrounds of stories could be powerful, unconstrained and imaginative, but their plots should be logical and common. More real the stories are, more immersed the players are, and more sensations of direct engagement are accepted. For example, the United Nations World Food Programme (WFP) developed an educational game named Food Force. Through the role playing of WFP staff, the players try to transport food to a virtual island damaged by wars, seek refugees, airdrop succor materials, fight with enemies and plan rebuilding programs for the farm. This kind of story background is novel and attractive, and the plot is practical. The players will assume themselves as heroes and take part in the tasks actively.

Different from our living environment, in the situation of game, it allows us to act according to our free aspirations despite of considering the results. Being like the learning environment, which should not restrain the capabilities of learners to construct their own knowledge structures, the game situation should not restrain the cognitive processes of players. It should allow players to try freely with their own aspirations, and choose when and where to start or stop game. Therefore, the plot in a story background should be designed as multi-choice, which will let the players try with curiosity, take part in their own choosing plot and enhance the sense of direct engagement as well as the enduring interest of the game.

4.6 High Intensity of Interaction and Feedback–To Consider the Interactive Feedback on Quantity, Accuracy and Cue-Sound

Interaction and feedback connecting with the objective dimensions are important elements to encourage the study activities. Communicating with the environment of games, the players will know whether their action is positive, within the rules and close

to the objectives according to the feedback information. A certain degree of interaction and feedback will guide the players to approach next objective and achieve the final victory.

However, when the feedback and interaction information is too many and exact, it will give negative influence to the game. For example, in the game called Day Off, the AI feedback system, Bob, will not only inform the players of the numbers and positions of the faults, which will block the further thinking, but its boring alarm sound will also disturb the concentration and game experience of players. Different modes of feedback will cause different feelings of players. The modes of feedback should be designed according to the different operations to the game. Therefore, a scientific consideration to the modes of feedback should be carried out on the quantity, accuracy and cue-sound.

4.7 A Continual Feeling of Challenge–To Provide the Skill-Balance Challenge through the Adjustment of Game Hardness

Challenge is the core element of the games, and the learner will strive to research the strategy to increase his ability to overcome the system when the challenge balances the skill. Therefore, the designer of educational game should set up different grades for different learners with different skills, which are selectable by the learner. At the same time, the hardness of game will be adjusted automatically according to the ability and performance of the learner. That is to say, there will be clear and reachable challenges step by step in accordance with the improvement of the learner's skills. And the hardness of challenge will be increased accordingly. The learner will continue the game in the sense of self-affirmation and self-fulfillment whenever he overcomes the challenge. That is the inner motivation of game.

4.8 Motivation–To Produce Motivation through Grade and Empirical Value

The game should reflect the impact of grades and empirical values on the learner, which means different grade and empirical value enjoy different power, which can force the learners to continue the game in the purpose of higher power. The designer can set up different grade according to different teaching objective to provide learning objective step by step to the learners; and the designer also can set up different empirical value according to sub-goal in the same grade. Take "Virtual Life" as instance, the participants can accumulate intellect, fascination and energy through different tasks, and when the empirical value reaches a certain extent, the participants can upgrade to a higher grade and to reach the complete learning objective finally.

5 Conclusion

The research and development of educational game in our country is still in a primary stage, effective models and ideas for educational game design are of great lack. This paper proposed the EFM model, in order to provide a new thought for researchers and developers of educational game. In the follow-up study, we will further study how to build the game learning environment, and take the model in educational game design practice to see how the model can help to improve it.

References

1. Keller, J.M.: Motivational Design of Instruction. In: Reigeluth, C.M. (ed.) Instructional Design Theories and Models: An Overview of Their Current Status. Erlbalum, Hillsdale (1983)
2. Chan, T.S., Ahern, T.C.: Targeting Motivation–Adapting Flow Theory to Instructional Design. Journal of Educational Computing Research 21(2), 151–163 (1999)
3. Kiili, K.: Evaluations of an Experiential Gaming Model. An Interdisciplinary Journal on Humans in ICT Environments 2(2), 187–201 (2006)
4. Wan, L., Zhao, M., Zhao, C.: Viewing the Design of Digital Educational Game from Experiential Games-Learning Model (in Chinese). China Audiovisual Education 10, 5–8 (2006)
5. Zhong, Z.: Discussing the Design of Learning Environment (in Chinese). E-education Research 7, 35–41 (2005)
6. Zhu, X.: Discussing the Design of Learning Environment (in Chinese). China Audiovisual Education 7, 16–18 (1996)
7. Houser, R., Deloach, S.: Learning from Games: Seven Principles of Effective Design. Technical Communication 45(3), 319–329 (1998)
8. Zhao, H., Zhu, Z.: The Analysis of Definitions and Typology about Educational Game (in Chinese). In: 10th GCCCE2006 Collected Papers, pp. 39–46. TsingHua University Publishing House, Beijing (2006)

Towards Generalised Accessibility of Computer Games

Dominique Archambault[1], Thomas Gaudy[2], Klaus Miesenberger[3], Stéphane Natkin[2], and Rolland Ossmann[3]

[1] Université Pierre et Marie Curie, INOVA/UFR 919,
9, quai Saint Bernard, 75252 Paris cedex 5, France,
[2] Centre de Recherche en Informatique du Cnam
292, rue St Martin, 75003 Paris, France
[3] Johannes Kepler Universität Linz, "Institut Integriert Studieren"
Altenbergerstrasse 69, A-4040 Linz, Austria

Abstract. Computer games accessibility have initially been regarded as an area of minor importance as there were much more "serious" topics to focus on. Today, the society is slowly moving forward in the direction of accessibility and the conditions come to make new proposals for mainstream game accessibility. In this paper we'll show the main reasons why it is necessary to progress in this direction, then we'll explain how works standard computer applications accessibility and why it is not working in general with games. We will discuss the state of the art in this area and finally we will introduce our vision of future accessibility framework allowing games developer to design accessible games as well as assistive providers the possibility of developing Assistive Games Interfaces.

1 Introduction

Computer games have become an important part in child and youth culture, and most children, in developed countries, have considerable experience of such games. Additionally these games are used by a growing part of the population, including especially young adults (on average 25 years old, including 40% of women[1]) but the proportion of players is also growing in other age groups of the population.

A lot of people with impairment are excluded from the computer games world because of accessibility. Indeed games accessibility have initially been regarded as an area of minor importance as there were much more "serious" topics to focus on. Since the middle of the nineties, lots of works have been focusing on making office computer applications accessible, and it's a fact that nowadays word processor and spreadsheets applications are reasonably accessible as well as web browser and mail readers.

Today, as Zyda claims, "*the time has come to take computer games seriously, really seriously*" [1]. Indeed the mainstream commercial market for computer games and other multimedia products has shown an impressive growth in the last five years. Costs for the development of a game may reach the level of major movie productions, involving more than a hundred employees [2]. The expectation by games players of ever

[1] TNS Sofres, Le marché français des jeux vidéo (The market of video games in France). afjv, November 2006. http://www.afjv.com/press0611/061122 marche jeux video france.htm

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 518–527, 2008.

more impressive games has seen increasing development budgets, and with a more focused use of new technologies.

Academia and also R&D over the last years have started to focus on "*serious games*". Leading experts speak of "*creating a science of games*" [1] with the goal of implementing games and game like interfaces of general importance for a growing number of applications and as a general trend in the design of Human-Computer Interfaces (HCI) [3].

In addition, general HCI is beginning to use concepts and methods derived from games as they promise an increased level of usability. Games and game-like interfaces [2]are recognised as a means to implement educational, training, general HCI and web applications with usability and effectiveness. Particular examples of interest are:

- eLearning and edutainment which more and more implement or use didactic games [4]. As an example of this trend it should be noted that critical issues like mathematics and science education are approached with game-based learning infrastructures and edutainment to address the well known didactic problems in this domain. "*Games stimulate chemical changes in the brain that promote learning.*" [5].
- Avatar based interfaces. We assist to a growing number of applications in such environments: for instance in France, real Job interviews have been organised in Second Life2 .
- Emerging Non Classical Interfaces (e.g. virtual/augmented reality, embedded systems, pervasive computing).
- A lot of Cultural Multimedia products, like Museum CD-Roms or DVD-Roms.
- Web 2.0
- Other software considered as inaccessible until today might come under accessibility discussions based on the principles, guidelines and tools developed for games and games like interfaces (e.g. simulation software, charts, virtual/ augmented reality).

Then, even if it would be considered as questionable to use the limited resources that are available for research on accessibility to address problems of people with disabilities using games or edutainment software, the general evolution of HCI towards game-like interfaces compel us to also consider a "serious" look at games from the accessibility perspective, in order to keep pace with the general level of accessibility achieved over the last decades in standard HCI. When standard HCI changes also accessibility has to change and this is closely related to games.

People with disabilities form one of those groups benefiting most from ICT[3]. Indeed, Assistive Technology (AT) enables them in a lot of situations in their daily lives, at school as well as at work or at home, in mobility, etc. The possibilities offered to the by eInclusion makes a difference in the life of a lot of people. Therefore it seems important that children get used to using technology as early as possible. Computer games are often a good training for the use of AT, for children as well as for adults after accidents, diseases. In addition playing games contributes considerably in establishing and ameliorating the skills in dealing with HCI.

---

[2] Linden Lab, http://www.secondlife.com

[3] Information and Communication Technologies.

From a different perspective, new approaches towards therapeutic and educational games for people with disabilities; children can benefit a lot from the use of computer games for their psycho motor and cognitive development [6].

We can now find a few hundreds of specific games, which have been developed especially for various groups of disabled users, but actually:

- this number is very short regarding to mainstream games, and they are often limited to one language,
- these games are usually very specifically dedicated to a extremely small group of end users with little or no access to the mainstream market (based on their abilities)
- these games are often very simple or old fashioned (even if a few very interesting exceptions exist)
- an important amount of these games are driven by specific pedagogical and therapeutic objectives and, on the whole, not much fun.

The limited budgets dedicated to specific developments make it very difficult to propose specific games with the quality and the size of mainstream games, which limits the possibilities of gaming experience for those players. Because of this, games for people with disabilities tend to worsen the segregation of disabled people from the mainstream gaming community they are the only games that they can interact with. This situation is in contradiction with the general eInclusive principles of ICT and Assistive Technology.

Accessibility of games is a more complex problem than software or web accessibility in general. The first reason, which seems obvious but is very important, is that: Accessible games must still be games! [7] Designing games that work for players with disabilities is quite a challenge: an important research, practical and social issue that has to be carried out now. This research should lead to one goal: the accessibility of mainstream games. Several aspects have to be taken into account: to find out how to handle game interaction situations with alternative devices, to develop models allowing to make mainstream games compatible with these alternative devices, to write according guidelines, methodologies and techniques.

To give people with disabilities the chance to have access to multimedia games should be seen as a great challenge for better eInclusion and participation in society.

The main groups of people addressed by these accessibility issues are those who cannot use mainstream games because their disability prevent them to use a modality which is necessary for some kind of games, namely:

- People who cannot use the ordinary graphical interface, because they are totally blind or because they have a severe visual impairment (sight rated < 0.05) [8] ;
- People who cannot use or have limited access to ordinary input devices like keyboard, mouse, joystick or game pad due to limited hand dexterity ;
- People with cognitive problems who need support to better understand the scene and react properly (e.g. symbol, text, speech and easy to understand support) ;
- People with hearing problems or deafness not able to accommodate to sound based interaction modalities ;
- People with problems in reacting to a strict time setting of the game out of various functional, cognitive and also psychological problems.

2 Software Accessibility

Today it is state of the art that people with disabilities can interact with the standard desktop/WIMP[4]based interface using Assistive Technology. Specific access software applications, like screen readers and screen magnifiers, alternative input devices, alternative information rendering – sound, text, signs, colour/size/contrast of objects, allow them to access to many computer applications. This is mainly the case, as mentioned above, for text based software: word processors, spreadsheets, mail clients, web browsers. The problem is that these access software applications are not able to access any software application whatever the way it has been developed. Indeed they need to collect accessible information from the applications to render it using alternative output modalities, or to control them using alternative input modalities.

In other terms, to achieve accessibility of software applications, it is necessary to have accessibility support embedded in the applications. During the last decade, accessibility frameworks have been developed and are available in the main environments. For instance, Microsoft has developed Microsoft Active Accessibility[5], to make their Windows Applications accessible, application developers have to implement the IAccessible interface[6]. There exist similar frameworks on Mac[7] and on Linux desktop environments[8]. Theoretical works can be cited too [9]. Furthermore specific development frameworks need to support accessibility, for instance Java[9] and Mozilla[10] .

It is not enough that applications respect accessibility standards. In most cases, content must be accessible too. For instance, in the case of a web site, the accessibility of web browser is necessary but the web contents must also be accessible. Graphical elements for instance must have textual alternatives, and this depends on the content itself. In that respect, the W3C launched the Web Accessibility Initiative to developed Web Content Accessibility Guidelines [10]. These guidelines indicate how to use each of the HTML tags to make a web site accessible. Accessibility of content has been developed in other content formats such as the proprietary PDF and Flash formats[11].

Of course there are still a lot of barriers in access to software and web content, but basically there are technical solutions asking for according political and practical measures to put this potential in place.

3 What Is Different in the Case of Games?

These accessibility solutions are working satisfactorily for standard desktop applications (WIMP based) but not for computer games. First the notion of working satisfactorily is a) not enough and b) not easy to define in that context.

---

[4] WIMP: Windows/Menus/Icons/Pointers
[5] MSAA:http://msdn2.microsoft.com/en-us/library/ms697707.asp
[6] http://msdn2.microsoft.com/en-us/library/accessibility.iaccessible.aspx
[7] Apple accessibility: http://www.apple.com/accessbility
[8] Gnome Accessibility: http://developer.gnome.org/projects/gap
KDE Accessibility: http://httpaccessbility.kde.org
[9] Desktop Java Accessibility
[10] Mozilla Accessibility Project: http://www.mozilla.org/access
[11] Adobe Accessibility Resource Center: http://www.adobe.com/accessibility

Indeed, the results of a game can not be easily quantified, like in the standard case of classical desktop applications. In a word processing software, it is easy to measure the time needed by a user to write a document or to edit a document produced by a colleague. In a game we can as well observe if a player succeeds, and measure the time to finish a level or any case relevant for the game considered. But this is far not enough. Unlike others software, games have to provide special good feelings to players. There are probably some emotional factors to consider in the desktop applications, but they are usually not taken into account, or at least unless they affect the productivity. In the case of a game these factors are the most important.

Video images and audio messages contain emotional components and specific patterns which can easily be perceived and due to empathy can be experienced by the viewer and listener. For the same reason interactive video games are attractive and popular among the youth and adolescence. Empathic arousal has a strong influence on people viewing, listening and reading by forming their social response to the external events through mental estimation of the problem and simulation of possible solutions and actions [11, 12].

It has been demonstrated in numerous psychological studies that some emotions can motivate a specific human action and behaviour. The development of the emotional intelligence in youth depends on a social inclusion and personal experience which usually rely on observing others' actions and behaviours presented in a real life (the cultural milieu) and in the artificial situations disseminated by movies, television and video games [13–16]. Being deprived of access to information with emotionally rich content, blind and visually impaired children have experienced a significant emotional distress which can lead to depression and deceleration in cognitive development [17, 18].

As we stated above: Accessible games must still be games! Visually impaired adults in work situation accept relatively big constraint on usability to be able to use the same software as their sighted workmates and to work on the same documents. This is not the case with children, especially playing. In other terms it is not enough to find a technical way allowing to access to all information needed in the interface, the result must be as interesting and as usable as the original game, and additionally it must be possible to succeed!

This helps us to understand that game interfaces are of a profound different nature than standard HCI and use their own technology (game engines). Usability and accessibility ask for freedom in time, speed, undo, mode of interaction,... It is a key criterion outlined by the W3C/WAI guidelines and software accessibility guidelines that the interface must not prescribe a certain interaction behaviour. But it is the core idea of games for realising immersion into a game, joy and setting up the gaming feeling to prescribe a restricted action and reaction behaviour and to force the user to be successful in this "reality". It seems that the more the player has to follow a strict behaviour, the more it seems that the game "takes the player into it" and puts immersion in place.

Therefore game accessibility goes beyond standard HCI and content accessibility measures. It must allow the prescription of behaviour by the system but it asks for alternatives and freedom of adaptation in the level of prescription and usage of modalities of interaction. If a mainstream game has put accessibility in place it is the role of adapted AT interfaces to realise immersion. Therefore it is inevitable to work on these adapted AT interfaces in game accessibility.

4 Game Accessibility during the Last Decade

Even if a fair amount of games has been developed in this field during the last 5 years, today there are still very few games that are accessible (including specific games and mainstream games). Coming back to the year 2000, one could only find a very short number of games usable by disabled players.

4.1 Specific Games

The first period that we identified is the period 2000-2005, that we will call the "basic studies". During this period we have seen the development of various games specifically designed for specific groups of people with disabilities. These kind of games are usually funded by foundations or non-profit organisations. Most of them are very nice for the group they were developed for but have little interest for mainstream, except maybe a few of the audio games. What is additionally of importance in the context of this proposal is that they demonstrate how to render various interaction situations with alternative modalities. This can be supplemented by the number of research papers about new uses of various modalities in the game play (Braille devices, haptic...).

The largest number of such games concerns audio games. Actually audio games include three different concepts in which the main play modality is audio.

The first meaning involve mainstream video rhythm games like Guitar Hero II. The second definition is related to artistic musical experiments. The third correspond to games which is based on sound environment (sound scenes, sound characters, actions) and can be played without vision, and therefore are accessible to visually impaired players (like interactive audio books, stories and tales): In 10 years, over 400 accessible audio games have been developed (which is very small as compared to video games).

The web site http://audiogame.net refers interesting interactive works. There exist a few visual audio games which can be very impressive and playable as well with or without sight. Terraformers [19] was developed with accessibility as part of the original concept of the game. On the other hand, AudioQuake [20] was developed as a research project to make a non-accessible game accessible.

A few tactile games can be found, that are games where the inputs and/or the outputs are done by tactile boards or by Braille displays, in combination with usually audio feedback. The use of Braille displays for gaming is only experimental by now. Some research is currently carried out in order to find models to represent a 2D space on a linear Braille display [21]. A few experimental games were designed in order to evaluate these models, for instance a snake game and a maze game. During the TiM project [22] (IST-2000-25298), a number of tactile games have been created or adapted from existing mainstream contents. Tomteboda resource centre in Sweden have published a report where they try to stimulate parents and educators to develop their own games using tactile boards [23].

[24] studied the possibilities offered by haptic technologies for creating new interactions usable by blind people. He worked especially with the Senseable Phantom. Then a number of papers explore the possibilities of using haptics in experimental games: [25–30].

The outcomes of this period is that we can base our work now on a lot of studies on playing games in various situations of functional limitation and on the adaptation of computer game situations.

4.2 Setting up the Foundations

The second period, ongoing since 2005, sees the emergence of the notion of games that work for all. This is declined in 2 aspects: game designed for all and accessibility of mainstream games. It is driven by the already mentioned fact that games and game like interfaces are recognised as important contributions to the next generation of HCI, eLearning and other applications.

The goal of games being developed under the title of designed for all is to give players with all different kinds of abilities or disabilities the opportunity to play these games. This requires a very advanced game setting and configuration. UA-Chess [31] is a universally accessible Internet-based chess game that can be concurrently played by two gamers with different (dis)abilities, using a variety of alternative input/output modalities and techniques in any combination. Access Invaders [32] is a designed for all implementation of the famous computer game Space Invaders, with the target groups of people with hand-motor impairments, blind people, people with deteriorated vision, people with mild memory/cognitive impairments and novice players. The approach of [33] was to make an already published game accessible and demonstrate the feasibility and the effort necessary to fulfil this goal. It is based on an open-source implementation of Tic-Tac-Toe.

Games designed for all must be seen as examples of good practise, demonstrating that Universal Access is a challenge and not utopia. In this projects we have to admit that the various alternative access features to these games require more development than the rest of the game itself.

Following these experiments it became clear that the accessibility of mainstream computer games needed to be improved. [7] proposes a set of rules to make computer games accessible by visually impaired people, derived from the TiM project. We started to work on formulating Guidelines for the Development of Accessible Computer Games, covering a wide range of disability groups [34]. IGDA[12] published a white paper on Accessibility [35], showing early signs of interest from the mainstream gaming industry.

5 What Is Needed Now?

We have developed the reasons why computer game accessibility should be considered seriously. Then we have seen how accessibility works in the case of standard desktop applications and why current accessibility frameworks it would not work with games or game-like applications. In the previous section we have seen that a lot of works have been studying how to render different game situations using different kinds of alternative devices.

Players with disabilities need to use Assistive Technology to play accessible games. But contrarily to any other computer application, this must not take off the characteristics of these applications that make them games. It is not only the task which one fulfils with an application (e.g. with office/mail software) but it is the procedure of

[12] International Game Developers Association.

playing the game it self which is fun and which provides learning benefits. In other terms, games accessed with AT still must be games and due to this it challenges the usage of AT. Then increasing accessibility of games will mean developing a new generation of assistive software taking into account much more parameters than current AT has access to, via the existing accessibility frameworks: characterisation of information available (including ranking of importance regarding the current task to fulfil), relative importance of events, speed, etc.

These new assistive software applications, which we will call Assistive Game Interfaces (AGI) will not likely be unique for a specific kind of impairment (like today a screen reader allows to access any office application). Depending of the ability constraints, some could be dedicated to a specific game or a game engine, some would be dedicated to a kind of games and finally some others would be generic (covering a large range of games).

We could for instance imagine a ”captioning application” allowing lots of different games to have captions when a character is speaking. On the other hand, for blind gamers, we could have specific AGI for text based games, another AGI working with a specific game engine, and a third one dedicated to a popular car race (since in this case the interaction would have to be completely redesigned).

To achieve these goals, the AGI will need to collect information from the core of the game itself. Indeed most of the information needed cannot be efficiently processed automatically from the mainstream game (for instance the captioning information). We have seen that the existing accessibility frameworks are not sufficient to provide these AGI with the necessary information. This means that it is necessary to design a new Game Accessibility Framework (GAF). This framework will have to take into account the specific data needed by various alternative devices to work properly. To continue with the example of the Captioning application, this application will need access to the complete transcription of the texts spoken by the characters in the game. The Game Accessibility Framework will have to specify how and in what format.

The first steps to carry out the specification of the Game Accessibility Framework are (a) a typology of game interaction situations and (b) a characterisation of Accessibility in terms of functional requirements. From the study of these expected results, the specification of the GAF can be produced, including the data formats and exchange protocols to transmit information between game and AGI.

Now it is time to make a significant move. This implies some participation from assistive technology specialists as well as from mainstream games developers.

The proposed solution may seem not realistic but it has to be considered that:

- The state of the art shows that the technology is ready
- This solution will be the lighter for game developers (consider for instance the work that would be needed to add a ”caption” option in a game, compared to the implementation of the access to texts that are already existing somewhere in the production process).
- the societal need for improving inclusion is growing in some leading countries (Northern European countries, Austria, Canada, Japan, etc) and the political pressure will necessarily follow, leading to laws and recommendations. We expect that this situation extends to the rest of Europe and North America, and to the rest of the world.
- the evolution of standard HCI towards game like interfaces will soon make these applications enter in the scope of existing laws

References

[1] Zyda, M.: Creating a science of games. ACM Communications 50(7) (July 2007)

[2] Natkin, S.: Video games and Interactive Media: A Glimpse at New Digital Entertainment. AK Peters (2006)

[3] Kellogg, W., Ellis, J., Thomas, J.: Towards supple enterprises: Learning from N64's Super Mario 64, Wii Bowling, and a Corporate Second Life. In: "Supple Interfaces": Designing and evaluating for richer human connections and experiences (September 2007)

[4] Chatham, R.E.: Games for training. ACM Communications 50(7) (July 2007)

[5] Mayo, M.: Games for science and engineering education. ACM Communications 50(7) (July (2007)

[6] Hildén, A., Svensson, H.: Can All Young Disabled Children Play at the Computer. In: Miesenberger, K., Klaus, J., Zagler, W. (eds.) ICCHP 2002. LNCS, vol. 2398. Springer, Heidelberg (2002)

[7] Archambault, D., Olivier, D., Svensson, H.: Computer games that work for visually impaired children. In: Stephanidis, C. (ed.) Proceedings of HCI International 2005 Conference (11th International Conference on Human-Computer Interaction), Las Vegas, Nevada, July 2005, 8 pages (proceedings on CD-Rom) (2005)

[8] Buaud, A., Svensson, H., Archambault, D., Burger, D.: Multimedia games for visually impaired children. In: Miesenberger, K., Klaus, J., Zagler, W. (eds.) ICCHP 2002. LNCS, vol. 2398, pp. 173–180. Springer, Heidelberg (2002)

[9] van Hees, K., Engelen, J.: Non-visual access to guis: Leveraging abstract user interfaces. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 1063–1070. Springer, Heidelberg (2006)

[10] W3C: Web Accessibility Initiative — Web Content Accessibility Guidelines 1.0. Technical report, World Wide Web Consortium (W3C) (May 1999), `http://www.w3.org/TR/WAI-WEBCONTENT`

[11] Grézes, J., Decety, J.: Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. Human Brain Mapping 12, 1–19 (2001)

[12] Prinz, W., Meltzoff, A.: An introduction to the imitative mind and brain. In: Meltzoff, A., Prinz, W. (eds.) The imitative mind: Development, evolution and brain bases, pp. 1–15. University Press, Cambridge (2002)

[13] Segall, M.H., Campbell, D.T., Herskovits, M.J.: The influence of culture on visual perception. Studies in Art Education 10(1), 68–71 (1968)

[14] Patterson, J.: Theoretical secrets for intelligent software. Theory into Practice 22(4), 267–271 (1983)

[15] Fromme, J.: Computer Games as a Part of Children's Culture. The International Journal of Computer Game Research 3(1) (May 2003)

[16] Sebanz, N., Knoblich, G., Prinz, W.: Representing others' actions: just like one's own? Cognition 88, B11–B21 (2003)

[17] Barresi, J., Moore, C.: Intentional relations and social understanding. Behavioral & Brain Sciences 19(1), 107–122 (1996)

[18] Chartrand, T.L., Bargh, J.A.: The chameleon effect: the perception-behavior link and social interaction. Journal of Personality and Social Psychology 76, 893–910 (1999)

[19] Westin, T.: Game accessibility case study: Terraformers -a real-time 3d graphic game. In: Proceedings of the Fifth International Conference on Disability, Virtual Reality and Associated Technologies, Oxford, UK, pp. 95–100 (2004)

[20] Atkinson, M.T., Gucukoglu, S., Machin, C.H.C., Lawrence, A.E.: Making the mainstream accessible: What's in a game? In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 380–387. Springer, Heidelberg (2006)
[21] Sepchat, A., Monmarché, N., Slimane, M., Archambault, D.: Semi automatic generator of tactile video games for visually impaired children. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 372–379. Springer, Heidelberg (2006)
[22] Archambault, D.: The TiM Project: Overview of Results. In: Miesenberger, K., Klaus, J., Zagler, W., Burger, D. (eds.) ICCHP 2004. LNCS, vol. 3118, pp. 248–256. Springer, Heidelberg (2004)
[23] Hammarlund, J.: Computer play for children who are severelly visually impaired: Using an alternative keyboard with tactile overlays. Technical Report 20, Tomteboda resource centre, Stockholm, Sweden (1999)
[24] Sjöström, C.: The sense of touch provides new computer interaction techniques for disabled people. Technology and Disability 10(1), 45–52 (1999)
[25] Johansson, A.J., Linde, J.: Using Simple Force Feedback Mechanisms as Haptic Visualization Tools. In: Proc. of the 16th IEEE Instrumentation and Measurement Technology Conference (1999)
[26] Wang, Q., Levesque, V., Pasquero, J., Hayward, V.: A Haptic Memory Game using the STRESS2 Tactile Display. In: CHI2006, Montréal, Québec, Canada. ACM, New York (April 2006)
[27] Raisamo, R., Patomäki, S., Hasu, M., Pasto, V.: Design and evaluation of a tactile memory game for visually impaired children. Interacting with Computers 19(2), 196–205 (2007)
[28] Evreinov, G., Evreinova, T., Raisamo, R.: Mobile games for training tactile perception. In: Rauterberg, M. (ed.) ICEC 2004. LNCS, vol. 3166, pp. 468–475. Springer, Heidelberg (2004)
[29] Crossan, A., Brewster, S.: Two-handed navigation in a haptic virtual environment. In: CHI2006, Montréal, Québec, Canada. ACM, New York (extended abstracts, 2006)
[30] Rodet, X., Lambert, J.P., Cahen, R., Gaudy, T., Guedy, F., Gosselin, F., Mobuchon, P.: Study of haptic and visual interaction for sound and music control in the Phase project. In: Proceedings of the 2005 conference on New interfaces for musical expression, Vancourer, pp. 109–114 (May 2005)
[31] Grammenos, D., Savidis, A., Stephanidis, C.: Ua-chess: A universally accessible board game. In: Salvendy, G. (ed.) Proceedings of the 3rd International Conference on Universal Access in Human-Computer Interaction, Las Vegas, Nevada (July 2005)
[32] Grammenos, D., Savidis, A., Georgalis, Y., Stephanidis, C.: Access invaders: Developing a universally accessible action game. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 388–395. Springer, Heidelberg (2006)
[33] Ossmann, R., Archambault, D., Miesenberger, K.: Computer game accessibility: From specific games to accessible games. In: Mehdi, Q., Mtenzi, F., Duggan, B., McAtamney, H. (eds.) Proceedings of CGAMES 2006 Conference (9th International Conference on Computer Games), Dublin, Ireland, pp. 104–108 (November 2006)
[34] Tollefsen, M., Flyen, A.: Internet and accessible entertainment. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 396–402. Springer, Heidelberg (2006)
[35] International Game Developers Association: Accessibility in games: Motivations and approaches (2004), `http://www.igda.org/accessibility/IGDA_Accessibility_WhitePaper.pdf`

Designing Narratology-Based Educational Games with Non-players

Yavuz Inal, Turkan Karakus, and Kursat Cagiltay

Computer Education and Instructional Technology,
Middle East Technical University, Ankara, Turkey

Abstract. Challenges of designing an educational game cause an ongoing debate that while one side proposes ludology as the key for a computer game, other side proposes narratology as the most important part of game environment. Ludologic attributes of games have been preferred more than narrative ones. However, results of studies attempted to reveal importance of narrative structures and storytelling for computer games, especially for educational ones. In the present study, narratology including storytelling and narrative structures will be discussed in terms of narratology-based educational computer game design. 46 non-player preservice teachers at Department of Foreign Language Education participated to the study. Participants, as subject matter experts on teaching English Language, designed educational game prototypes. Those prototypes were analyzed and reported according to their narrative aspects.

Keywords: Narratology, educational game, game design, non-players, ludology.

1 Introduction

Games are basic human activities and they are older than culture (Huizinga, 2006). Computer games are also driven from social, cultural and economical aspects of societies. They are inseparable part of entertainment especially for youngsters (Fromme, 2003). Squire (2002) reports that games can be used to explore historical issues, investigate some complex learning situations, and they can manage and govern some digital places even whole civilization. Squire and Jenkins (2003) state educational values of games like socialization, interaction and understanding the concepts which are aimed to give players are important for learning.

Games are investigated according to two major aspects; ludology and narratology (Ang, 2006). These two aspects define complexity and interactivity of a game environment. Especially narratives seem to be important issue for interactivity of player and the game emotionally and narrative structures of games provide complexity of games (Lindley, 2002). Characteristics of interaction in educational games make students active, decision maker and problem solver. Narratives of games seem play an important role in educational games for cognitively engaging students during learning process. Therefore, in design processes of games, narrative structures and storytelling parts should be constructed properly.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 528–534, 2008.

1.1 Importance of Narratives

Narratives provide decision points and interaction with game environment. Embedding an educational context within a story and narrative structure might be beneficial and easy for game designers during game design process. Therefore, it seems that narratives are as important parts as ludology parts in educational games. Mateas and Stern (2005) stated that small pieces of narratives provide a response from players sequentially in a dynamic system, thus narratives can provide a local and global agency to provide players valuable experiences. According to Mallon and Webb (2005), narratives are not simple beings which can be analyzed in end-unit level, rather they are complex beings which need detailed "specifications, properties and concrete examples" to show the narratives in games.

According to Squire and Jenkins (2003) story of component of games is important part of educational games especially for science education. They especially emphasized that stories of games should be explored in a way that stories help imagination of students and what can be done for students to learn with stories in games. They also added some research results that stories influence decisions of students related their future career. Mallon and Webb (2005) urged that conflicts between game, interaction, hyper-structure and narratives should be solved and empirical studies should be conducted to reveal several solutions between gameplay and narrative structure. They also argued that lack of empirical studies that put what aspects of games provide the significant experiences for players.

1.2 Game Design and Narratives

Computer games, especially role playing ones, make students force to follow and think on problems (Squire & Jenkins, 2003). Narratives in games are important in role playing games because narratives include who the player is, what he is supposed to do, which decisions should be made etc. In this case, it should be considered that narratives in games should be designed wisely to guide students' learning process throughout gameplay. Reiber (2001) stated that children like story of games, challenge, competition and the subject matters in game environment while playing educational games.

Juul (2001) believes that in games, players also produce stories, and games include some narrative parts. Players can realize these narrative elements during the play, and games and narratives have similar common structures. Therefore, in educational games, developers should take into account integration of content with stories and make students actively participate in these stories. Besides, teachers' contributions are invaluable for computer games (Squire & Jenkins, 2003) because teachers know about target groups and how context and content can be integrated in a compatible way. In this study, researchers' aim was to understand how preservice foreign language teachers embed educational context within a game environment and how they organize storytelling and narrative structure in the game.

2 Methodology

In this study, the following research problems were examined;

1) What are the roles / missions / duties of avatars in designed educational game environment?
2) What are the major patterns of preservice teachers in terms of storytelling in educational games?
3) What type of narrative structures do preservice teachers prefer to embed within the educational game environment?
4) What types of game genres do preservice teachers prefer to design educational games?

2.1 Data Collection

Totally 46 preservice teachers at the Department of Foreign Language Education participated to the study. A demographic questionnaire to figure out students' computer game habits, preferences, reasons if they do not play computer games, places in which they play computer games, and their demographic characteristics from Durdu, Tufekci and Cagiltay (2005) were administered. Besides, a second questionnaire including open-ended questions concerning game prototype scenarios of 46 preservice teachers was employed.

2.2 Overall Demographics of Preservice Teachers

Participants, as subject matter experts on English Language teaching, designed educational game prototypes basing on narratology in games. Majority of the participants were female (n=37) and rests of them were male (n=9). Almost noneof them were serious game player.

37 of the participants had computer at their homes. 18 of them stated that they use computer more than 10 hours, 14 of them use between 5-10 hours and rest of them use up to 5 hours per week. Although participants in the study preferred using computers, 40 of them stated that they do not like to play computer games. Only 6 of the participants stated that they play games from time to time. The games that they prefer are mostly Solitaire like games, quiz-trivia games, puzzle games and racing games. The reasons why participants didn't prefer playing were respectively "having no time to play", "don't know how to play", "not interested in playing" and "games are time consuming".

3 Results

In the study, students were asked to design educational games to teach English as a second language basing on narratology. Results were categorized as "subject matter", "avatars in game environment", "patterns of storytelling", and "designed game genre for educational games".

3.1 Subject Matter

In the study, students were asked to select a subject matter concerning the scope of English Teaching. Majority of them (n=24) preferred teaching "vocabulary" in a game-based learning environment. Teaching names of the animals, numbers, countries, and shopping terms were most preferred topics under the vocabulary teaching. "Directions" were second favorite subject matter (n=10) that students preferred. Finding directions to save someone or reach target aimed in an adventure in game environment were considered as a subject. Besides, teaching "grammar" was other favored subject matter. Totally, 8 students stated that game-based learning environments might be appropriate for teaching grammar.

3.2 Avatars in the Game Environment

All of the students in the study preferred avatar-based educational game throughout their designs. They gave a role or mission to avatars and story of the game was based on these avatars. One of the students stated that "There is a pilot who encounters a master caution during the flight. Therefore, he needs help immediately in order to tend, so he tries to get help from the airspace".

Most of the students tried to embed educational context in narrative structures and storytelling in games, and avatars are the vital part of their game designs. They asked players immerse avatars during game play, and while doing missions of avatars, players could learn the topic easily. For instance, one of the students stated "There is a sick child coughing seriously, and saying "please help me!". The mission of the player is to help the child gathering clues in the game environment concerning the illness. The game includes four stages and the player will try to find all illness by giving right answers of vocabularies in the environment". One of the interesting findings of the study was that almost none of the students specified their avatars as male or female. They only determined the missions or roles of the avatars player will use, but they didn't prefer explaining gender of avatars.

3.3 Patterns of Storytelling

Four storytelling patterns were categorized as "finding a missing stuff" such as an object, avatar, animal etc in game environment, "helping an avatar in game environment", "finding directions" in a labyrinth, puzzle or city to reach the given target, and "trying to manipulate or design the game environment".

Most of the students preferred to design a game environment to find some missing stuff. They aimed to make player try to find and answers of questions related to educational context. One of the students stated "10 animals escaped from city zoo. Authorities tired to do their best to find the missing animals but some of them are still somewhere in the city. Player has to find the animals but first s/he has to find clues by giving correct answers of the questions". Mainly, in order to teach vocabulary or grammar, students preferred such storytelling to present answer/question environment for players. A female student stated "Mother and her child went for shopping during a weekend. While they were shopping, the child got lost suddenly and the mother was in panic. She asked information desk to find her child and they assigned an employee

to solve this problem. In the game, player has to find the child". To students, designing educational games behind such storytelling patterns was more educational because player has to try to find correct answers or questions related to educational context. Therefore, so as to motivate students and keep their attention high, before asking question, such stories might be helpful.

All of the students preferred to use avatars while designing an educational game and they preferred to give a mission or role to players. One of these missions was helping an avatar (a character) in the game environment. Students tried to base stories by saving human-beings in the world from disasters or catastrophes. One of the students stated "Some of the animal species have been disappearing suddenly in the ecology. However, this disappearing is not results of deaths of such animals. It was suspected that some guys have been kidnapping these animals. The mission of the player is following clues related to countries and their characteristics and give correct answers to questions so as to find kidnappers". Another interesting storytelling patterns in this category was growing an object or bringing up a baby in the game environment. Students aimed to teach vocabulary concerning an object or baby while dealing with.

Finding directions in a labyrinth or puzzle environment was one of the most favorite storytelling patterns of students. For the game environment, generally jungle, city, under the sea or zoo was preferred. One of the students stated "In order to find the treasure under the sea, the one thing that player has to do was to give correct answers of animal names and meaning of them when they were asked. After each correct answer, the player will take a clue and it will help him/her to reach the treasure." Similarly, another student stated "A group of students was taken to city zoo to visit the animals. But, because the zoo was extremely large, some of them got lost. According to the map they have, they will try to find exit of the zoo".

Some of the students, especially females, preferred to design game environment which are similar to Sim City or Barby-like games. They aimed to teach grammar or vocabulary to players while designing game environment or dressing avatars with clothes or other accessories. A female student stated "There is a 3D avatar in the game environment. Players will select their own avatar at the initial point of the game. Then, they will dress selected avatars in a shopping center by getting correct answer of meaning of the clothes or accessories."

3.4 Designed Game Genre for Educational Games

Majority of the students preferred adventure games. According to Dickey (2006), storytelling and narrative structures play a prominent role in such game design since "Adventure games are often considered as a form of interactive fiction" (Dillon, 2003). According to the results of the study, avatar-based educational games were preferred by the students. They gave prominent role to avatars to teach an educational context. Avatars had a mission or role and players had to give correct answers to reach determined target or finish the game. The game genres that students preferred were categorized as respectively, adventure games, quiz-trivia games and Barbie-like games.

4 Discussion

In the present study, preservice teachers' educational game preferences basing on storytelling and narrative structures in terms of teaching an educational context concerning foreign language teaching were investigated. Results of the study revealed that avatar-based educational games were mostly preferred. Participants aimed to provide story by using an avatar and explain educational purposes with avatar-based game design.

For vocabulary learning, there are many games and teachers prefer to create a game for vocabulary. Role-playing games are most preferred one because most of the preservice teachers give a mission to players. In this study, preservice teachers gave importance to vocabulary teaching. They stated that a role-playing game basing an avatar-based design might have beneficial and educationally valuable for learners. Johnson, Vilhjalmsson and Marsella (2005) give importance to mission games in language education because in these kinds of games, player has several target actions, players are engaged to complete their missions, and they should interact with other avatars by using verbal communication, and attempt to make some tasks. In foreign language education, conversation is important to make practice with grammar and vocabulary. Thus, it seems that selection of game type was related to teaching methods of foreign language in classrooms.

In the study stories of games were also related to real life. The participants preferred some stories which might happen in classroom, helping someone to find direction, stating the names of some objects and making conversations with some people. This might be caused from teachers' game playing habits and game preferences because participants of the present study were almost non-players. Dickey (2005) argues that playing games provide new insights to develop different games. In this context because teachers have little experience with games, it might be said that they depended on their teaching strategies while designing game. Besides, teachers tended to develop detailed stories rather giving just rules and actions in the game that was because stories are important elements for games according to the participants. With this implication it can be assumed that stories are attractive elements for foreign language teachers to design / develop / or prefer educational games.

5 Implications

Computer game applications have been emerged in classroom settings for educational purposes. However, it has been investigated for several years that effective and efficient educational game design issues are not clear for both educators and game designers. There need to be conducted studies focusing on game design basing on educators' preferences and needs. In the present study, a gap on educational game design issue from educators' points of views will be aimed to fulfill.

With the results of the study, it might be concluded that narratology-based game design might be appropriate for educational games that are related to foreign language teaching. Especially designing avatar-based educational games was very popular among participants of the study. Developing a game by giving a mission to players might be a way while designing a game. Besides, when the educational subjects that

participants preferred were considered, mostly vocabulary teaching was most preferred among all. We believe tat results of the present study will guide to both educators and educational game designers while designing games with foreign language teaching.

References

1. Ang, C.S.: Rules, gameplay, and narratives in video games. Simulation & Gaming 37(3), 306–325 (2006)
2. Dickey, M.D.: Engaging By Design: How Engagement Strategies in Popular Computer and Video Games Can Inform Instructional Design. Educational Technology Research and Development 53(2), 67–83 (2005)
3. Durdu, P.O., Tufekci, A., Cagiltay, K.: A comperative study between METU and Gazi University students: Game playing characteristics and Game preferences of University Students. Euroasian Journal of Educational Research 19, 66–76 (2005)
4. Fromme, J.: Computer Games as a Part of Children's Culture. Game Studies 3(1) (2003), [online journal], (viewed 28 December 2007), http://www.gamestudies.org/0301/fromme/
5. Huizinga, J.: Homo Ludens. A Study of the play element in culture. Beacon Press (2006)
6. Johnson, W.L., Vilhjalmsson, H., Marsella, S.: Serious Games for Language Learning: How Much Game, How Much AI? In: Proceedings of AIED 2005:The 12th International Conference on Artificial Intelligence in Education, pp. 306–313. IOS Press, Amsterdam (2005)
7. Juul, J.: Games Telling stories? A brief note on games and narrative. GameStudies 2(1) (2001), [online journal] (viewed December 25, 2007), http://gamestudies.org/0101/juul-gts/
8. Lindley, C.A.: The gameplay Gestalt, narrative and interactive storytelling. In: Proceedings of computer games and digital cultures conference, June 6-8, 2002, Tampere, Finland (2002)
9. Mallon, B., Webb, B.: Stand Up and Take Your Place: Identifying Narrative Elements in Narrative Adventure and Role-Play Games. ACM Computers in Entertainment 3(1) (January-March 2005)
10. Mateas, M., Stern, A.: Build It to Understand It: Ludology Meets Narratology in Game Design Space. In: DIGRA Changing Views & Worlds in Play, Vancouver, June 17 (2005)
11. Squire, K., Jenkins, H.: Harnessing the Power of Games in Education. Insight, The Institute for the Advancement of Emerging Technology in Education 3(1), 5–33 (2003)
12. Squire, K.: Cultural framing of computer/video games. GameStudies 2(1), [online journal] (viewed December 28, 2007), http://www.gamestudies.org/0102/squire/

Interactive Game Development with a Projector-Camera System

Andy Ju An Wang

School of Computing and Software Engineering, Southern Polytechnic State University
1100 South Marietta Parkway, Marietta, GA 30060, USA
jwang@spsu.edu

Abstract. This paper reports our experience on interactive game development with a projector-camera system in a special topic course. We describe the course objective, learning outcomes, and how we offered it in our curriculum. Ideas of the rationale and expansion of this course is also presented. A camera-projector system requires a digital camera, a projector, and a computer. The projector and camera could be installed in various locations depending on the application. Our system used a rear-installed (behind the player) projector and camera, with players standing in between the projector-camera pair and the projected surface – a screen or a wall. We discuss in this paper a simple testing game engine using OpenCV and DirectX, with a few example games developed by the student teams.

Keywords: Computing and programming, Entertainment computing, Computer games, Computer science education.

1 Introduction

The computing discipline is currently facing an unprecedented array of pressures to change. The enrollments in traditional Computer Science programs have been fluctuating for the last a few years while the skill set of computing professionals keeps growing in a rapid and unpredictable fashion. The computing profession itself is becoming more complex, with traditional disciplinary boundaries blurring or disappearing in several emerging technological areas. According to a recent study conducted by the Computing Research Association (CRA), the number of newly-declared computer science majors in the fall of 2007 was half of what it was in the fall of 2000 – 7,915 versus 15,958. On the other hand, students have many misconceptions about computing. Many believe that there are no opportunities and most do not see the connection between computing and innovations in fields such as entertainment. Some reforms in computing education have been undertaken in response to these challenges.

More and more video game-related courses and degree programs are being offered in colleges around the county in response to the digital media industry's need for skilled workers and the tastes of a new generation of students raised on Game Boy and Xbox [6] – [13]. In the spring semester of 2006, the author offered a special topic course entitled "IT 4903/6903 Entertainment Computing and Technology" at the School of

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 535–543, 2008.

Computing and Software Engineering, Southern Polytechnic State University. This course introduced students to the breadth and depth of the issues involved in the field of entertainment computing and technology. It discussed the background, concepts, technologies, impacts, and business models of entertainment computing, one of the most promising and exciting future computing areas. Students were exposed with different views and different techniques in this field. This was a project-based course where students were required to work in small teams. There was also a field trip to a local entertainment technology company and three of the five student teams were actually working on small projects sponsored by the company. This course was designed for senior students majored in Computer Science, Software Engineering, or Information Technology. One of the purposes of this course was to introduce to students an exciting field in entertainment computing, helping students begin to see computing as a rich source of educational and career opportunities. On completion of this course, students should be able to:

- Understand basic principles and techniques in entertainment computing.
- Exercise the skills needed to create entertainment computing applications.
- Demonstrate an understanding of the general programming concepts and methods for creating interactive entertainment applications.
- Evaluate and identify current artistic and commercial trends as related to the field of entertainment computing and technology. As a part of general education, this course should also help students to
- Communicate (written and verbally) about a complex, technical topic simply and coherently.
- Work and interact collaboratively in groups to examine, understand and explain key aspects of entertainment computing.

Entertainment takes an important role in our life by refreshing our mind and inspiring our creativity. With the rapid advancement of computer hardware and software, new forms of entertainment have emerged, such as video games, entertainment robots, and network games. Virtually every household has computing devices such as computers, televisions, entertainment robots, etc. [14]. Entertainment computing and technology bring both promises of enriched experience and risks of negative social impacts. Traditional video games do not allow players to have large freedom to interact with the game content physically other than through conventional I/O devices such as keyboard, mouse, or joysticks. The camera-projector-based games, however, allow players to control and interact with the game content through their physical movement and gestures. Moreover, the camera-projector-based games can be played any place where the game content can be projected onto a surface. For example, these games could be played in a classroom, gym, exercise room, or even on your lawn in your back yard. These kinds of "sweaty games" [15] have great educational value for college students as well.

In this paper, we briefly report how we organized and implemented an IT course covering camera-projector-based interactive game development and what had the students done in this course. The course combined camera-projector technology with traditional computer game design and implementation. The rest of the paper is organized as follows: Section 2 introduces an interactive game involved student-professor connection developed by one student team in the course. Section 3 introduces two

interactive games using the similar technology of camera-projector systems. Section 4 presents a preliminary prototype of a game engine designed for interactive games using the camera-projector system. Section 5 discusses related work and concludes the paper.

2 POP the Professor

One student team was assigned to develop an interactive game for new freshmen to learn more about their professors. An additional goal of this project was to make it a recruiting tool at the event of open house, attracting prospective students to computing disciplines, especially attracting more female students by showing images of female professors in our school. If prospective female students see that there are a significant number of female role models at the school they may be more likely to select one of the computing disciplines as their majors, such as Computer Science, Software Engineering, or Information Technology. The Initial requirements for the game were for the player to be actively involved in searching and capturing the cascading images of IT professors at SPSU. A conference room was also present in the middle of the playing field as depicted in Figure 1 below.

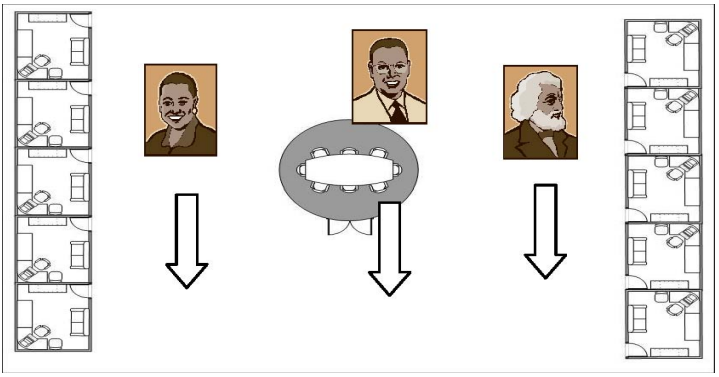


Fig. 1. The initial design of "Pop the Professor"

We used a camera-projector system in this game development. The images were projected on a wall, and the players stood between the projector and the wall, playing the game with their movement and gesture. As the images falls down into the playing field the player must make "contact" with the images and direct them to their office until the professor is in the correct office. Some detailed requirements include the following:

- If the correct office is found and the professor entered, the image occupies the office and the short biography of the professor is displayed with a popup window. A sound clip of the professor's voice is played to identify her or him.
- Once in the office, if the player's shadow falls on the desk top, a list of courses the professor teaches will be displayed with a popup.
- If the shadow falls on the couch or conference table then the professor's office hours are displayed with a popup.

- If the player attempts to place a male professor in the wrong office, and the office is a female professor's office, a scream is heard.
- If the player attempts to place a male professor in the wrong office, and the office is a male professor's office, a male voice rebukes the player.
- Images of the professor will be displayed to reflect complementary characteristics of the respective professor.
- Several images of students will also be in play. After the professor is in their own office the student can enter to consult with the professor. Students cannot enter a professor's office unless the professor is present. If the student attempts to enter an office without a professor, a sound clip will play.
- At a random time a sound clip announces a faculty meeting is to be held and all professors are to be herded into the central conference room.
- Professor's images are not recycled. When they drop to the bottom of the play field they gather there until they are struck or popped up by the players. The game is over when all professors are in their right office place. A congratulatory sound clip is played at the conclusion of the game.

Pop the Professor was based on PlayMotion's Critters game [1], which is a game that contains areas into which the player moves an appropriate game piece. The player moves a game piece with their hands and movement to match the appropriate place on the projected screen or on a wall. These locations on the wall correspond to an area called a grabber slot that matches the entity. A grabber slot can be moved to different places on the playing field. Pop the Professor required multiple grabbers that could be located on either side of the screen representing professor's offices as shown in Figure 2 below. Electronically generated sketches of each professor's face were presented to the player to allow them to associate each professor with the appropriate office. Individual cubes representing professors are released onto the playing field, the player then maneuvers the cube into the appropriate office area.

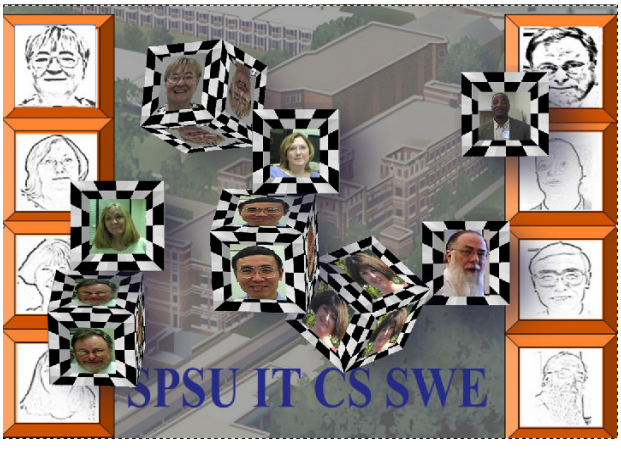


Fig. 2. "Pop the Professor" screenshot: Game begining

Upon success, i.e., placing all professors into their appropriate offices, a new screen is displayed as Figure 3 below, where the offices are moved to a different location. Appropriate sounds are displayed when the cube is hit and upon successful completion of the game, the professor's voice stating her/his name, etc.

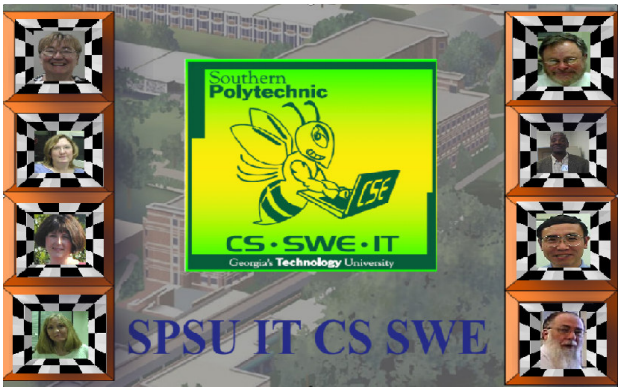


Fig. 3. "Pop the Professor" screenshot: Success

The feedback from students in developing this interactive game was very positive. There were a few significant learning points worth for a discussion here.

- The team was exposed to a unique interactive technology, the camera-projector system, in which the users of the system need no keyboard or mouse to play the game. Instead, the users play the game by standing before a wall screen and manipulate the game objects with their hands, body and movement.
- The team was exposed to a use of XML different from that seen in the usual Web design context. The ability to set attributes for program elements using XML is very powerful. In addition, it allows a level of abstraction such that the designer can configure the system characteristics without having to manipulate the actual game engine.
- Considerable knowledge was gained in capturing and manipulating audio and video files. This involved working with multiple open source applications. It allowed the team to evaluate and draw usability conclusions for many applications. The team discovered capabilities of Microsoft PowerPoint that they had not previously been exposed to. Another interesting achievement was the ability to create a background image for a page or presentation which looks similar to a watermark.

3 Drumline and College Life

The second project team developed an interactive game called "Drumline", using the camera-projector system. This game was designed to be visually educational, and allow users of all ages to interact with drum images that will be projected on any flat surface with our camera-projector system. The users could beat different kinds of drums that

would come across the screen, hearing the drum sound and learning from a pop-up text panel containing pertinent information about the culture of the country accompanied by a sound clip. A single user could play multiple drums while the drums were still moving across the displaying surface, either a wall or a screen. Multiple users could interact and play with the Drumline at the same time. After playing this game, the users would have enjoyed, understood, and learned more about drum-related cultures. Drumline included seven different kinds of drums from different continents. An XML file was used to capture the characteristics of a drum and could be configured easily. A sample XML file is shown below:

```
<?xml version="1.0" ?>
<TITLE
   baseDirec-
   tory="c:/Drumline/">
   <SOUNDTRACK volum="0"
   baseDirectory=".">

   <PATH
   file="IntroDjembedrum32.wa
   v"/>
   </SOUNDTRACK>
   <CONFIG>
        <BALL radius="1.2" mass="300"
        popUpDirectory="../Drumlinefacts/AfricanCongaDrum"
        popUpColor="1.3,1.0,2" tiltAngle="22.40"
        popUpSound="./Africandrum.wav" tex-
        ture="./AfricandrumCongaDrum.jpg"
        quantity="1" />
        <!--More drums defined here-->
        <TABLE
             forceAmplifier="0.8"
             width="32 height="24"
             catchTimeForPopUp="2500"
             popupTimeout="5000"
             shadowTimeout="5"
             kineticFriction="0"
             staticFriction="0"
             backgroundFile="../../data/Drumline2/5DjembW.jpg"
             gravityMagnitude="0.03"
             collisionDetect="tight"
             roll="on"
             popUpSound="../../data/sounds/popup_open.wav"
             popDownSound="../../data/sounds/popup_close.wav"
             clackSound="../../data/sounds/Whomp-soft.wav"
             screenWrapping="on"
             shadow="../../data/textures/ballshadow.tga"
             glowSprite="../../data/space/glow.tga" />
   </CONFIG>
</TITLE>
```

The third project team developed an interactive game called “College Life”, targeting new college students in a way that they could learn how to make a smooth transition from high school life to college life, balancing their life and activities in a college atmosphere. Activities were represented by images falling from the top of the

screen. The players could grab objects that would have consequences on their everyday life and future achievement. A score system was designed to challenge the player and encourage them to make the right decisions. For instance, if the user grabbed SPSU 101 Orientation logo, her score would be increased by 50 points for participating in the hitchhike course. The user had many objects to choose from that represent different activities and social activities, some of which include parties, attending classes, studying, tests, exercises, work, and dating. Selecting too many of either category of activities would reduce the points, but selecting a well-balanced activities would make the score increase until the graduation ball dropped then the game was over.

4 A Testing Game Engine

A camera-projector system requires a digital camera, a projector, and a computer. The projector could be installed in the back of players or mounted on the ceiling. The camera, however, could be installed anywhere including in front of, in rear of, or on top of players (mounted on the ceiling). Ceiling installation of a heavy projector is difficult and not portable. For simplicity, we use rear installing projector and camera as illustrated in Figure 4 below:

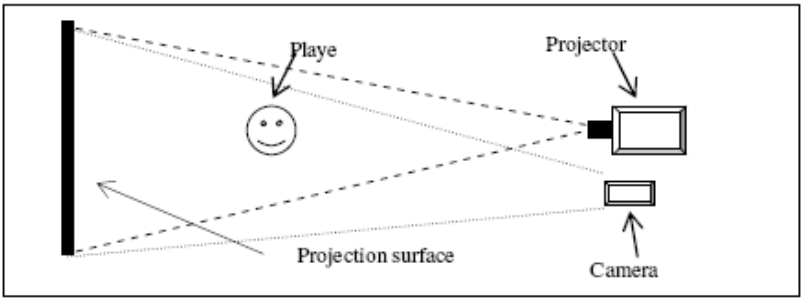


Fig. 4. The Camera-projector system

Our experiment utilized OpenCV and DirectX. OpenCV was used for image/vision processing and DirectX taking the result and implementing it to the game. OpenCV is an open source library for computer vision and image processing. It provides support for applications such as human-computer interaction (HCI), object identification, segmentation and recognitions face recognition, gesture recognition, motion tracking, ego motion, motion understanding, and mobile robotics [2]. DirectX [3] is a Microsoft product used especially in game programming because it contains numbers of Application Programming Interface (API) to help users develop game. Components of DirectX that are extremely useful are DirectDraw (produce raster graphics), Direct3D (for 3D graphics), DirectSound (for playback and recording sound), DirectInput (process data from keyboard, mouse, joystick, or game controllers), DirectMusic (for playback soundtracks), DirectSetup (for DirectX components installation), and DirectX Media (for animation).

The initial purpose of this project was to create an interactive game where players can have fun and physical exercise at the same time when playing the game. The integration required that the game provide both proper content physical excise and game play factors for enjoyment. Our game used a camera-projector system as depicted in Figure 4, where the camera is used to capture the player motion and gesture image and map it to the game content by OpenCV. The objects are drawn using DirectX component (DirectDraw) and used as the "balls" in a typical "brick and balls" game. The Mouse pointer and Bar are also created using DirectDraw. Background music is used because the nature of this game is motion action and the combination of the two would give the player more entertainment experience. There was a configuration file where all the features of the game could be modified including the bar, objects, mouse pointer, background image, background music, and click sound. This student team built this interactive game using the camera-projector system we discussed in previous section. From design to implementation and testing, a prototype was implemented. The game idea was not easy to put into realization without a solid teamwork and significant effort. Throughout the project, students learned to use some tools that they were not familiar with such as OpenCV and DirectX programming. Other tools that were useful for the project include Audacity, which is a powerful tool for sound editing and a freeware, GIMP, which is also a freeware similar in its functions to Photoshop, and lastly Microsoft PowerPoint, which can be a practical tool for creating background and it is easy to use.

5 Discussion and Conclusion

According to [21], about one-third of US students intending to major in engineering switch majors before graduating. In comparison, 38% of all undergraduates in South Korea receive their degrees in natural science or engineering. In France, the figure is 47%, in China, 50%, and in Singapore 67%. In the US, only 15% of graduates receive their degrees in engineering. For students, there are multiple causes for the decline in interest in the computing professions. Students perceive opportunities in computing as rapidly vanishing; the misconception and misleading image of computing professionals; and a poor early education leaves many unprepared for and disinclined to the computing fields. The decline in students in computing is clearly more serious among women and minorities. The author believes that a multiple solution strategy must be adopted to change this trend. One of the solutions is to establish specific programs which challenge and attract students to study and work in computing field. Through the practice of offering IT 4903/6903 Entertainment Computing and Technology course, the author saw clear evidence that students would become more engaged in computing if we enhance computing curriculum with game-related sophisticated environments and more tangible results to coding problems. With the positive feedback from students about this course, the department is currently considering to offer more game-related courses and environments.

It has been an active research area in combining projection technology with computer vision and computer gaming in the last a few years. Example applications include display walls, interactive display surfaces, intelligent environments and performance

art. The author is working on the combination of smart mobile devices with the camera-projector system to provide more immersive and rich experience in interactive games.

References

[1] PlayMotion, http://www.playmotion.com
[2] OpenCV Wiki. (Retrieved April 26, 2006), http://opencvlibrary.sourceforge.net headab972f644b0ef223ecaf1b3d7c17d6458c7cb0e5
[3] Pike, A.: DirectX 8 Tutorial (Retrieved March 15 2006), http://www.andypike.com/tutorials/directx8/
[4] Flood, M.: C++ Config File Library (2003) (Retrieved April 18, 2006), http://rudeserver.com/config/index.html
[5] McDaniel, T.L.: Intel Open Source Computer Vision Library Version 4.0-Beta, Installation and Getting Started Guide for Windows [Electronic Version] (Retrieved March 20, 2006)
[6] B.S. in Computer Games Development, DePaul University, http://www.cs.depaul.edu/programs/2005/BachelorGAM2005.asp
[7] WPI Interactive Media and Game Development course: http://www.wpi.edu/Academics/Majors/IMGD/Academics/imgdcourses.html
[8] CMU Entertainment Technology Center: http://www.etc.cmu.edu/
[9] USC Entertainment Technology Center: http://www.etcenter.org/
[10] The Academy of Game Entertainment Technology: http://www.academyofget.com/
[11] Entertainment Law, http://academy.smc.edu/curriculum/coursedescriptions.html
[12] Entertainment minor: http://snow.sierranevada.edu/~csci/eteksnc.html
[13] Academy of Entertainment & Technology, Santa Monica College, Santa Monica, CA 90405, http://academy.smc.edu/
[14] Greene, K.: Intel Inside Living Rooms, MIT Technology Review, December 14 (2005)
[15] Associate Press, Sweaty Video Games, MIT Technology Review, November 17 (2005)
[16] Wilson, A., Oliver, N.: Multimodal Sensing for Explicit and Implicit Interaction, Microsoft Research, Redmond, WA
[17] Wilson, A.: TouchLight: An Imaging Touch Screen and Display for Gesture-Based Interaction. In: ICMI 2004, State College, Pennsylvania, USA, October 13–15 (2004)
[18] Wilson, A.: PlayAnywhere: A Compact Interactive Tabletop Projection-Vision System, Microsoft Research (2005)
[19] Sukthankar, R., et al.: Smarter Presentations: Exploiting Homography in Camera-Projector Systems. In: Proceedings of International Conference on Computer Vision (2001)
[20] Rhyne, T.-M.: Computer Games' Influence on Scientific and Information, IEEE Computer (December 2000)
[21] Boylan, M.: Assessing Changes in Student Interest in Engineering Careers Over the Last Decade, CASEE, National Academy of Engineering (2004)
[22] Microsoft Research White Paper: Computer Gaming to Enhance CS Curriculum (2006)

Animated Impostors Manipulation for Real-Time Display in Games Design

Youwei Yuan[1] and Lamei Yan[2]

[1] School of Computer & Software, Hangzhou Dianzi University
(Hangzhou, 310018, China)
[2] School of Printing Engineering, Hangzhou Dianzi University
(Hangzhou, 310018, China)
y.yw@163.com

Abstract. This paper describes a system platform for animated impostors manipulation for real-time display in games design. Our "agent common environment" provides built-in commands for perception and for acting, while the in-between step of reasoning and behavior computation is defined through an external, extendible, and parameterized collection of behavioral plug-ins. Finally we introduce concrete case studies that demonstrate the effectiveness of our approach.

Keywords: Animated; level-of-detail(LOD); impostors; image caching; image-based rendering.

1 Introduction

In 3D computer graphics, an additional consideration is that rendering complex shapes realistically requires significant resources, when very often the same shapes can be effectively depicted in a line drawing style that uses less data, modeling effort, and computation time. The most basic non-photorealistic rendering uses little or no shading, and simply draws lines along silhouettes and sharp features. Recently, Decarlo et al. introduced suggestive contours [1]. These are additional view-dependent lines (like silhouettes) that convey a more complete impression of shape while adding relatively few additional lines. Rendering is then equivalent to a resampling process where surfels are blended with a Gaussian distribution in the image space[2].

The computational workload in graphics processing systems is generally split between a central processing unit (CPU) and a graphics processing unit (GPU). A combination of software, firmware and/or hardware may be used to implement graphics processing. For example, graphics processing, including rendering can be carried out in a graphics card, graphics subsystem, graphics processor, graphics or rendering pipeline, and/or a graphics application programming interface (API), such as OpenGL.

The vertex shader is traditionally used to perform vertex transformations along with per-vertex computations[3].Once the rasterizer has converted the transformed primitives to pixels, the pixel shader can compute each fragment's color. This pipeline

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 544–550, 2008.

is further extended in the upcoming generation of DirectXr10 hardware, introducing an additional programmable geometry shader stage. This stage accepts vertices generated by the vertex shader as input and, unlike the previous stage, has access to the entire primitive information as well as its adjacency information.

In recent years, there has been a dramatic increase in the processing power of GPUs, which are now typically able to distribute rendering computations over a number of parallel hardware pipelines. This has led to the transition of several stages of the rendering pipeline from the CPU to one or more GPUs. It is beneficial then, to make the most efficient use of the computational abilities in both the CPU and the GPU. Any increases in efficiency can be directly translated to increased realism and speed, while also reducing cost.

In our method, the 3D model will be rendered from these views, and the set of renders will be stored into an image texture. This texture is usually organized as a grid, where columns share views from the same slice and rows share views of the same stack of the bounding sphere. In the case of animated models, a discrete set of frames will also be selected, and for each frame, a set of views will be rendered and stored. Images produced by this preprocessing stage will be later used to obtain the texture map applied for rendering impostors in games design.

2 The System Structure

The system structure of animated impostors manipulation for real-time display in games design is shown in fig.1.

The core of the system understands a set of commands to control a simulation. A method for using a graphics processing unit (GPU) to cull an object database, comprising: (a) encoding per-object parameters in texture format thereby creating at least one per-object texture containing said encoded per-object parameters; (b) updating a fragment program on the GPU, said fragment program embodying a culling operation; (c) Create and place different virtual humans, objects, and smart objects (objects with interactivity information)r (d) produces cull results for a set of database objects, whereby said produced cull results will eliminate or reduce further processing of invisible, occluded, or distant objects. (f) Apply a motion motor to a virtual human. Examples of such motion motors are: key-frame animation, inverse kinematics [4], a walking motor[5], facial expressions, etc. These motors can be triggered in parallel and are correctly blended and composed, according to given priorities, by a specific internal module [6]. (g) Query pipelines of perception [7] for a given virtual human. Such pipelines can be configured in order to simulate, for example, a synthetic vision. In this case, the perception query will return a list with all objects perceived inside the specified range and field of view.

Trigger a smart object interaction with a virtual human. Each smart object keeps a list of its available interactions, which depends on the object internal state. Each interaction is described by simple plans that are pre-defined with the use of a specific graphical user interface. These plans describe the correct sequence of motion motors to accomplish an interaction. The GUI is used to interactively define the 3D parameters needed to initialize the motion motors, as positions to put the hand, movements to apply at object parts, etc.

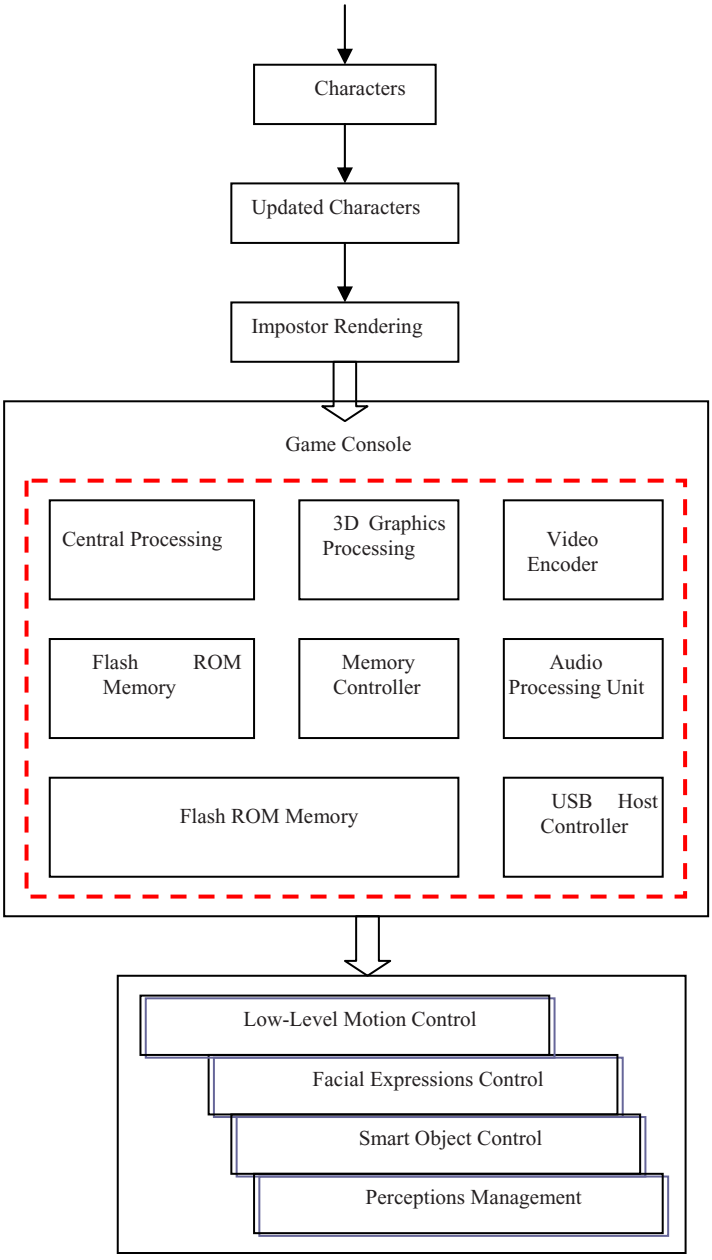


Fig. 1. The system structure of animated impostors manipulation for real-time display in games design

3 Multi-resolution Modeling

In this section we show how a multi-resolution virtual human can successfully be constructed and animated.

We first use Predefined Animations and Motion Capture: in many cases, the poses of animations will be known in advance. This may be the case either due to manual animation of frames, or the result of motion capture data. In the known examples, some bones may have imperceptible movement, and it may be visually acceptable to simplify their skin polygons as if they did not deform. From the examples given, we can compute the probability distribution of the configurations directly, and use this information to guide our simplification.

Our real-time virtual human model consists of an invisible skeleton and a skin. The underlying skeleton is a hierarchy of joints that correspond to the real human main joints. Each joint has a set of degrees of freedom, rotation and/or translation, which are constrained to authorized values based on real human mobility capabilities [8]. Unlike other attempts we did not model a multi-resolution skeleton because our purpose was not to demonstrate the effectiveness of animation level-of-detail. Hand joints can be replaced with a single joint though.

After carrying the surface along with the particles, it is deformed under the action of surface forces, similar to balloons. The forces are derived by minimizing the potential energy of the surface. The potential energy is composed of external potentials which depend on the particles and internal potentials which depend on the surfels. We derive an implicit and an attracting potential such that the energy is minimized when the surfels are attracted to an implicit surface and to the particles, respectively. Minimizing the internal potentials, consisting of the smoothing potential and the repulsion potential, yields a locally smooth and uniformly sampled surface. From the potential energy we derive forces acting on the surfels. While the derived forces from the implicit, attracting and smoothing potential act in normal direction, the repulsion force is applied in tangential direction.

Each primitive is attached to its proximal joint in the underlying human skeleton. The set of primitives then defines an implicit surface that approximates the real human skin. Sampling this implicit surface results in a polygonal mesh, which can be directly used for rendering. Sampling the implicit surface is done as follows: we start by defining contours circling around each limb link of the underlying skeleton. We then cast rays in a star-shaped manner for each contour, with ray origins sitting on the skeleton link. For each ray, we compute the outermost intersection point with the implicit surface surrounding the link. The intersection is a sample point on the cross-section contour.

Depth information is estimated from multiple fixed cameras and allows easy segmentation of the user from other people and background objects. An intensity-invariant color classifier detects regions of flesh tone on the user and is used to identify likely body part regions.

As for the head, hands and feet, we still have to rely on a traditional decimation technique to simplify the original mesh. Manual intervention is still needed at the end of this process to smooth the transition between LODs(Levels of details). The body

extremities can cleverly be replaced with simple textured geometry for the lowest resolution which dramatically cuts down the number of triangles.

Finally a face detection module is used to discriminate head regions from hands, legs, and other body parts. Knowledge of the location of the user's head in 3D is passed to the application. Motion model has a sequence of motion parameters for each action. In case of transmitted information with additional parameters such as 3-D positions, the inverse kinematics theory is used for generating motion parameters depending on the additional parameters. The system lets an avatar act according to a sequence of motion parameters corresponding to an action decided by behavior model. Fig.2 shown a few virtual humans in game design.

Fig. 2. Showed a few virtual humans in game design

4 Results

The proposed technique was evaluated on a Pentium Xeon computer at 3.2Ghz with 512 MB of memory, and rendered to a 1280 x 1024 window. A different number of characters was used to evaluate the performance of the proposed technique. Rendering performance for different number of characters is shown in figure3. Maximum frame rates were achieved when rendering all characters as impostors, while minimum frame rates involved a mixed rendering of impostors and instanced geometry. Table 1.shows different number of characters.

Fig.4 shown the interface of games we designed.

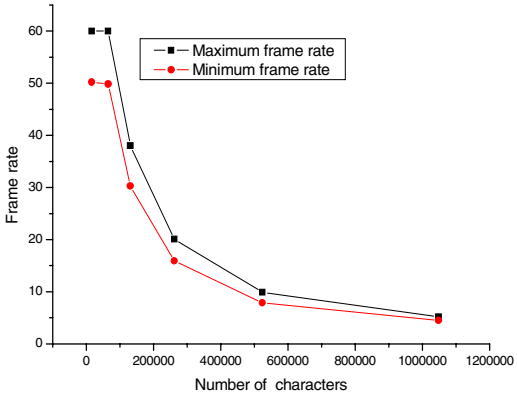


Fig. 3. Rendering performance for different number of characters

Table 1. Shows different number of characters

Number of characters	Minimum frame rate	Maximum frame rate
2^{14}-16384	50.2 fps	60.0 fps
2^{16}-65536	49.8 fps	60.0 fps
2^{17}-131072	30.3 fps	38.0 fps
2^{18}-262144	15.9 fps	20.1 fps
2^{19}-524288	7.9 fps	9.9 fps
2^{20}-1048576	4.5 fps	5.2 fps

Fig. 4. Shown the interface of games we designed

5 Conclusions and Future Work

Current design of the geometry shader stage can only generate individual primitives or lists via stream out. Once generated, there is no vertex reuse due to lack of associated index buffers for GPU-generated data. This affects performance for post-stream out rendering passes and triples required resulting vertex memory footprint.

An efficient technique has been presented to display large crowds of animated characters at game interactive frame rates. Impostors are a well suited technique for graphics processor, as a constant single quad is only required to display each character, and its animation and transformations are based on simple texture lookups. Here, as characters will be rendered to a larger on-screen area, the use of more instanced geometry will produce more realistic results.

Our method is based on impostors, a combination of traditional level-of-detail techniques and image-based rendering and relies on the principle of temporal coherence. It does not require special hardware (except texture mapping and Z-buffering capabilities, which are commonplace on high-end workstations nowadays) though fast texture paging and frame buffer texturing is desirable for optimal performance.

Finally, next generation graphics hardware should be thoroughly evaluated, as upcoming features, such as geometry shades, animated instancing, or vertex shades texture lookups, may provide with additional functionality that may be harnessed to

improve the flexibility and efficiency of GPU crowd rendering. Thus the hardware would be able to allocate appropriate storage for each invocation, as well as allocate the number of indices generated by this invocation.

References

1. DeCarlo, D., Finkelstein, A., Rusinkiewicz, S., Santella, A.: Suggestive contours for conveying shape. ACM Transactions on Graphics 22, 848–855 (2003)
2. Alexa, M., Behr, J., Cohenor, D., Fleishman, S., Levin, D., Silva, C.T.: Point Set Surfaces. In: Proc. of IEEE Visualization 2001, pp. 21–28 (2001)
3. Cabral, B., Cam, N., Foran, J.: Accelerated Volume Rendering and Tomographic Reconstruction using Texture Mapping Hardware. In: Proceedings of the 1994 Symposium on Volume Visualization, pp. 91–98 (1994)
4. Baerlocher, P., Boulic, R.: Task Priority Formulations for the Kinematic Control of Highly Redundant Articulated Structures. In: IEEE IROS 1998, Victoria (Canada), pp. 323–329 (1998)
5. Boulic, R., Magnenat-Thalmann, N., Thalmann, D.: A Global Human Walking Model with Real Time Kinematic Personification. The Visual Computer 6, 344–358 (1990)
6. Boulic, R., Becheiraz, P., Emering, L., Thalmann, D.: Integration of Motion Control Techniques for Virtual Human and Avatar Real-Time Animation. In: Proceedings of the VRST 1997, pp. 111–118 (1997)
7. Bordeux, C., Boulic, R., Thalmann, D.: An Efficient and Flexible Perception Pipeline for Autonomous Agents. In: Proceedings of Eurographics 1999, Milano, Italy, pp. 23–30 (1999)
8. Boulic., R., Capin, T.K., Huang, Z., Kalra., P., Lintermann., B., Magnenat-Thalmann, N., Moccozet, L., Molet, T., Pandzic, I., Saar, K., Schmitt, A., Shen, J., Thalmann, D.: The Humanoid Environment for Interactive Animation of Multiple Deformable Human Characters. In: Proceedings of Eurographics 1995, Maastricht, August 1995, pp. 337–348 (1995)

Virtual Avatar Enhanced Nonverbal Communication from Mobile Phones to PCs

Jiejie Zhu[1], Zhigeng Pan[1], Guilin Xu[1], Hongwei Yang[1], and David Adrian Cheok[2]

[1] State Key Lab of CAD and CG, Zhejiang University, Hangzhou, P.R. China
[2] Mixed Reality Lab, National University of Singapore, Singapore

Abstract. Nonverbal communication is a special kind of communication using wordless messages such as gesture, body language, posture, facial expression and eye contact. Such communications are specially attractive in virtual environments (VEs) which incorporating 3D avatars. Many of techniques for nonverbal communication in VEs have been studied and reported. However, transferring existing techniques to mobile platform are seldom reported. In this paper, we introduce our approach of creating a nonverbal communication environment between mobile phone and normal PCs. 3D face modeling is taken as an example to explain the system architecture. This modeling process is integrated with 3 platforms. The prior knowledge of modeling only uses one front view image which can be captured by built-in phone camera without high quality constrain. The two ends,between phone to phone or phone to PC, can download models from server and share the communication environment. Key techniques such as facial features detecting, face model personalizing are presented and experiment results show a lifelike face-to-face conversation can be simulated.

Keywords: 3G Phones, Mobile 3D, Radial Distortion, Facial Expression, Nonverbal Communication.

1 Introduction

Phones are built for human-to-human voice connection for long years. To bring people face-to-face in this situation, videophone technology are explored. These terminals connect voice and video stream which is captured by phone camera simultaneously. Thanks to recently growing bandwidth offered by wireless networks, such as UMTS, WLAN, WiMAX, and the success of H.264, acceptable real time videophone are available on mobile phones now.

Although mobile video phone has many advantages, no manipulation to the content may make troubles. Sometimes, people are unwilling to expose their real feelings or they may not show the real situation while talking to others. On the other side, most of people want to embed their digital representations, such as cartoon like characters, 3D talking heads and virtual avatars to mobile platform. At same time, people are likely to use mobile phone as an important information input to rich their digital repository.

The main contribution of this paper is providing techniques to enrich user's expressional content in their digital repository and studies technology transfer of content

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 551–561, 2008.

modification from mobile phones to PCs. We select mobile phones as an input device because it is most often used and most of them have built in cameras. We presents our work on facial feature extracting, 3D face personifying, facial expression generating, connecting technique both on mobile phone and Internet.

2 Related Work

A number of research of 3D face modeling have been proposed. Overall methods can be grouped into two parts: geometry-based and image-based[1]. Geometry based methods standardize the face framework which describes face model using grids, such as polygons, surfaces and volume grids. It requires special 3D scan equipment to get initialized data. The scanning process and data editing are tedious and time consuming. Image based methods avoid scanning by recovering 3D information from multiple views. Many vision information are extracted to accelerate this modeling process such as structure information, stereo correspondences, face shading region and face silhouettes. With prior knowledge of human face, image-based method achieves better appearance. However, they are subjective to image noise.

To improve the speed and quality of transmitting facial expression data on low-bite network, several facial animating specification are reported[2, 3]. The most widely used is Facial Action Coding System (FACS)[4]. A FACS coder "dissects" an observed expression, decomposing it into the specific Action Units (AUs) which produce the facial movement. The scores for a facial expression consist of the list of AUs. MPEG-4 face animation standard[5] is another widely used coding standard. It supports the transmission and composition of facial animation with natural video. Not like FACS, facial animation parameter (FAP) set is defined based on the study of minimal facial actions and is closely related to muscle actions. The FAP set enables model-based representation of natural or synthetic talking-head sequences and allows intelligible visual reproduction of facial expressions, emotions, and speech pronunciations at the receiver[6].

To accelerate the rendering speed of 3D models on mobile phones, 3D graphic hardware accelerators are designed and used. ARM[7] has implemented several types of graphic hardware accelerator for mobile phones. Based on their reports, the rendering capacity of ARM9, ARM10E and ARM11 series are all exceed the capabilities of PCs of 1995. To enhance compressing capacity, MPEG hardware accelerator is also embedded on mobile phone like SuperH Mobile Application Processor Series[8]. On-chip MPEG-4 hardware accelerator, it increases over twice MPEG-4 processing performance. SH-Mobile series handle most MPEG-4 processing by means of middleware, which make it possible to implement high-performance, low power consumption systems which incorporate moving-picture playback, Videophone, and similar sophisticated functions.

However, realistic 3D face modeling and efficiently coding on low bit-rate wireless network still remain an open issue. We focus on the first problem and presents our work. Since MPEG-4 hardware accelerator has been available on mobile phones, we choose MPEG-4 facial animation standard to code facial expression and transmit data both on wireless network and Internet. In addition, we use VRML to render 3D face models on Internet.

3 System Architecture: From Phone to Pcs

The conceptional idea of our system is to connect between phone to phone or phone to PC. Figure 1 illustrates our conceptional system architecture.

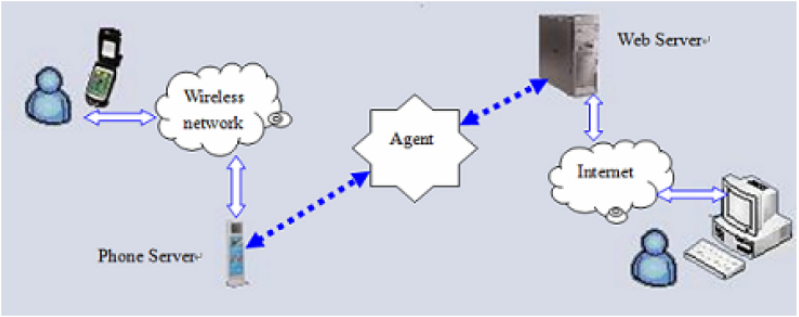


Fig. 1. Conceptional architecture with two types of terminals

To connect between mobile phones to PCs, the system architecture 3 platform. Figure 2 illustrates overall architecture with 3 different platforms.

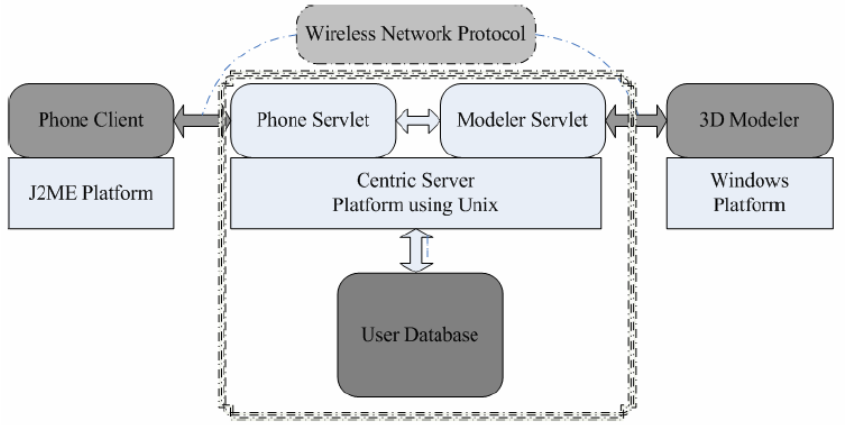


Fig. 2. System Architecture across 3 platforms: Mobile Client on J2ME; Centric Server on Unix; 3D Modeler on Windows

Phone Client: This client is presently regarded as mobile phones with built in cameras. To help user capture front view image and establish nonverbal communication, a user interface is designed and encapsulates internal mobile phone APIs. Two main sub-interfaces are implemented: facial feature point selection and personality customization. Without them, selecting facial features and changing their 3D accessories, such as hair color, skin color, configuring with glasses are unavailable. In addition, communicating scheme to centric server is implemented using HTTP protocol, which

is a sub-protocol of WAP. This connection makes upload, download and change resources available. All these mixed with voice connection, which means the client can talk while doing nonverbal communication at the same time.

Centric Server: This server takes a role as middleware between mobile phone and Internet. Each side is assigned one servlet to handle communications. On the client side, Phone Servlet is compatible with common mobile phone servers. On the Internet side, Modeler Servlet is used to exchange data between PCs and Centtric Server. While communicating, polling and pushing techniques make upload and download resources on this server available. The server also stores clients' information in a database which is accessible both by mobile phones and PCs.

3D Modeler: To reduce system burden on Centric Server, 3D face modeling is implemented on this PC by 6 steps (see Figure 3).

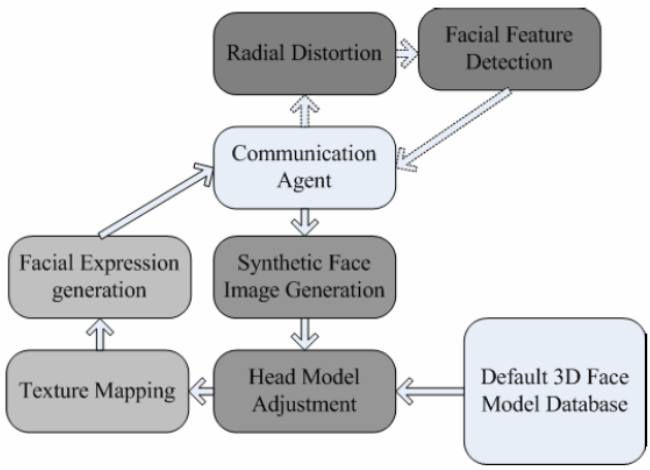


Fig. 3. Process of 3D face modeling

Radial Distortion removes image distortions; Facial Feature Detection detects facial features using Skin Color Possibility (SCP) and Gradient Change Map (GCM); Head Model Adjustment module morphs general model to a personalized one; Texture Mapping maps polished face image to the 3D face model and Facial Expression Generation creates several primary facial expressions by changing the position of detected facial features.

4 Key Techniques

4.1 User Interface of Mobile Client

A friendly interface is very important for mobile phone application. Figure 4 illustrates our implemented user interface.

Four layers are divided on the main screen from up to bottom. The up-most region shows the program's logo, memory size, battery power ; below this layer is a text

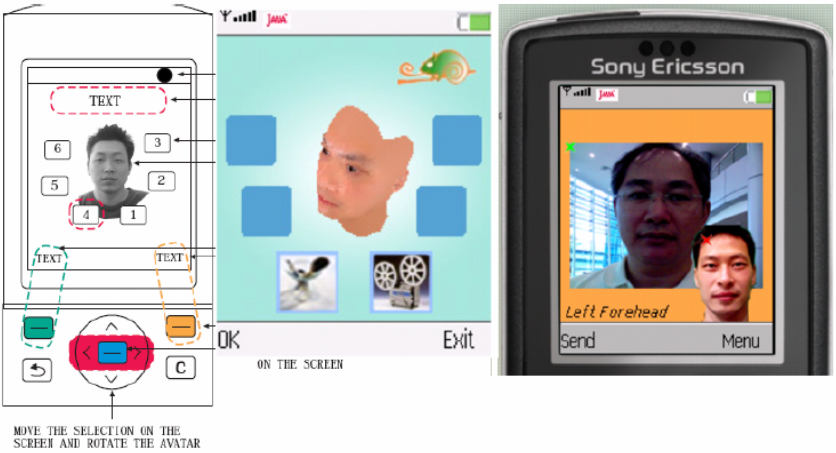


Fig. 4. User Interface of Mobile Client: Left Image is our concept of user interface; the center image is a given example of main menu; the right image is a given example of facial feature selection

highlights of selected functions; on the third layer, the 3D face model is rendered by JSR-184 functions[11] with 6 sub-menus surrounded it which rendered by MIDP2.0 APIs; on the last layer, it shows the description of function sequences.

4.2 Facial Features Detection

We use two steps to detect facial features on user's front view image. First step locates the face region by Skin Color Possibility (SCP). Second step detects facial features according to prior knowledge of human face's division.

To locate user's face region, we calculate the skin possibility of image pixels and match them with sample SCPs. The probability is clustered in a small area of the chromatic color space due to the color distribution of human skin and Gaussian distribution is used to calculate the likelihood of a skin pixel:

$$p(r,\mathrm{b}\,) = \exp[-0.5(x-m)^T C^{-1}(x-m)] \tag{1}$$

where, $x=(r,b)^T$; C is the covariance matrix of r to b , r and b is calculated by $r=r/(R+G+B)$) and $b=b/(R+G+G)$ respectively; m is a mean vector of r and b . Based on skin sample[15], we use double threshold method to detect skin pixels of front view face image.

To detect facial features, we first generate an Intensity Changing Map (ICM) by prior knowledge of human face (see Figure 5).

Both the horizon intensity change and vertical intensity changes are used to detect facial features. For example, eye region is first detected by the ratio in template. Its regional horizon intensity and vertical intensity changing values are then calculated. At last, the intersection of maximum horizon and vertical intensity changing point is selected as the eye centers. Figure 6 illustrates an example of this process.

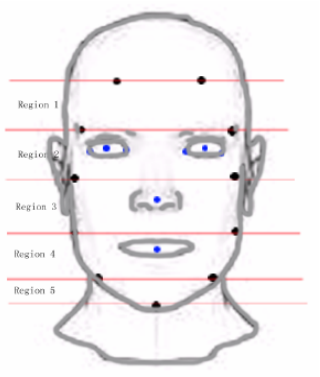


Fig. 5. Intensity changing map: 5 regions are divided on this face template by different ratios. In each region, intensity change is calculated and compared with templates.

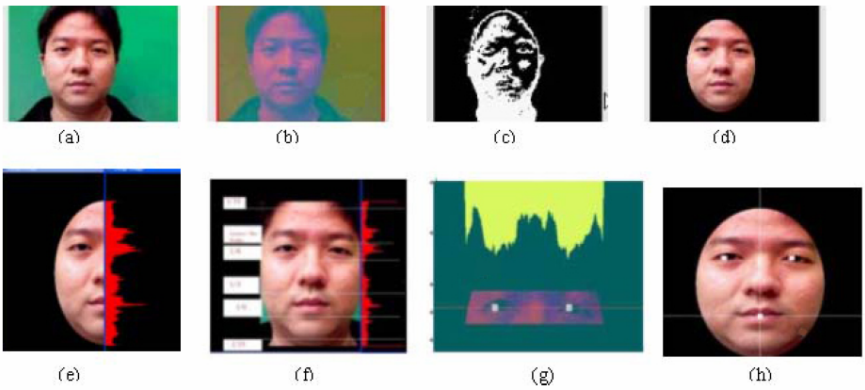


Fig. 6. Example Process of Eye and Mouth Centers Detection

(a) Original front view image. (b)Transforming RGB to YCC color space. (c)Recognizing face region using SCP and deleting isolated points by image morphology operation.

(d) Clipping the face region by ellipse fitting. (e)Two horizon gradient changes are calculated for half face. (f) Determining which feature by comparing with ICM. (g)Calculating vertical intensity changes. (h)Determining the facial feature points by maximum intensity intersections.

Considering the changes of illumination condition, the skin possibility may not correct. For example, the right side of the face in (c) is missed because the light is occluded by left side. So, users have to capture their face image under a global light environment. In addition, ellipse fitting requires high accuracy of isolated point deletion. Otherwise, the ellipse will be too large with background.

4.3 Face Texture Generation and Mapping

Before mapping the original face image to 3D face model, the image should be given extra lights and skin-like background colors to generate better visual results. We first

calculate the face region and this time using the stored facial features. Again, an ellipse approximation method is used to get the face region. Re-sampling algorithm resizes the image and image operations blend it to a predefined skin background image. Figure 7 shows a sample of the results.

In our system, cylinder mapping algorithm is used to map face texture to a 3D face model. Its process is shown in Figure

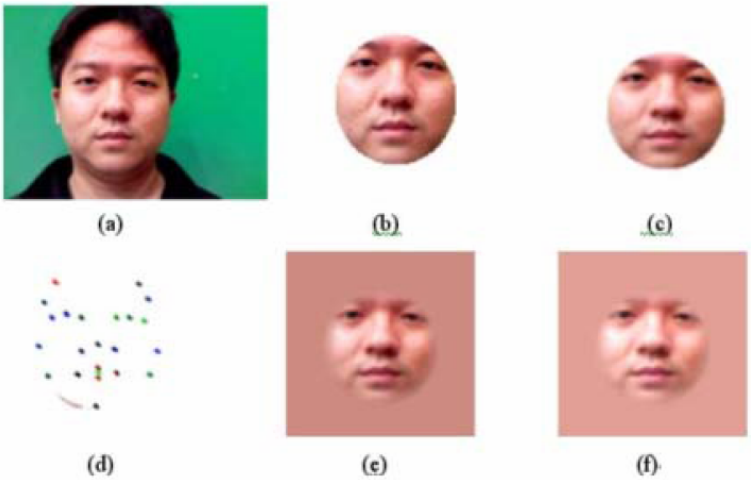


Fig. 7. Process of generating a visual face texture: (a)Original face image. (b)Ellipse face region. (c) Re-sampled face region. (d)Re-sampled facial features. (e)Blended to a skin background. (f)Final face texture after image operations.

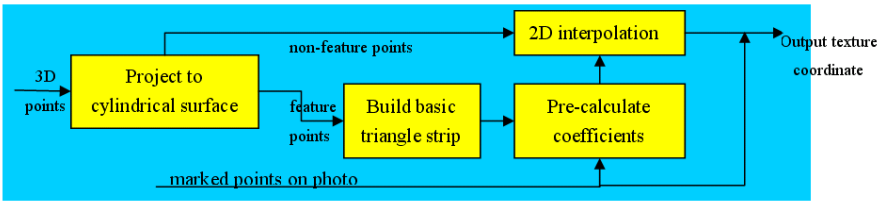


Fig. 8. Process of cylinder mapping

This process maps a facial points on 3D face model to facial points on 2D face texture. It first maps 3D points to a cylinder and then this cylinder is extended to an image plane. Hardware matching of texture coordinates is carried to all facial points, and non-facial points' texture coordinates is achieved using 2D interpolation. Before interpolation, the 3D points should build basic triangle strips. Since the number of 3D points are constant, we can predefine these triangular strips. Figure9 shows a solution using 23 points.

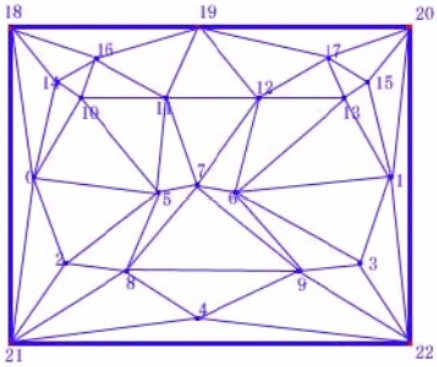


Fig. 9. Image triangular strips using 23 points

5 Facial Expression Generation

To generate facial expressions, our system provides 6 primary facial expressions: Joy, Sadness, Anger, Fear, Disgust and Surprise. Other expressions can be obtained while tracking the facial features from a video stream captured by built in camera of mobile phones. We assume that the 3D face model's mesh topology are all same, thus our algorithm can be applied to all 3D face models. To calculate the morphed facial points, a minimization equation of displacement on all facial points is designed:

$$E = \min \sum_i [d(x_i - x)^T C^{-1} d(x_i - x)] \tag{2}$$

where x_i is a short version of 3 direction of 3D points. (x_i, y_i, z_i) is the facial point's position and (x, y, z) is the referenced facial point. A least square method solves above equation with defined limitation of largest displacement range. From experiments results, weighted minimization of different axis gives better results. Non-facial features are interpolated by facial points and some of them are multi-weighted because they are influenced by different facial features. Figure10 shows two expressions created by high resolution texture and low resolution texture. Duce to small size of mobile phone's screen, low resolution texture in size of 160*120 is also acceptable for our application.

Figure 11 shows results of 6 facial expressions rendered on a real mobile phone. The mobile phone we used is Sony Ericsson Z800. Its screen size is 176*220, memory is 64M and the resolution of phone camera is 1.3 mega. The animation of each expression is quite good from feedbacks.

6 Discussion and Future Work

In this paper we introduced our work of creating a nonverbal communication environment across mobile phone to Internet. This is illustrated by a 3D face modeling across 3 different platforms.

However, based on our work, we find no standard Mobile3D APIs for all mobile phone available due to commercial protection. The most widely applied Mobile3D specification in embedded system is OpenGL ES, others like OpenKode, OpenML, Open VG and Microsoft's Direct3D Mobile are also used. In addition, different Mobile3D APIs are implemented using different language, like Mascot Capsule using C++ language and JSR-184 using Java language. All these prohibits a unified architecture design.

Low memory and small screen are another bottle necks for 3D application on mobile phone. The size of common 3G mobile phones' memory is 64M with 200M HZ of

Fig. 10. 2 examples of facial expressions on PC. Left is in high resolution, while right is in low resolution.

Fig. 11. 6 primary facial expressions rendered on a real mobile phone

processing speed. The highest terminals can reach 1G memory and 600 M Hz. The size of common 3G mobile phone's screen is 240320. Powered by batteries, rendering high quality of 3D graphic on these devices is challenging due to small resolution of the display. From research results[17], the average eye-to-pixel angle is larger for mobile phones than for PCs. This implies the quality in each pixel should be better for a mobile device than for a PC.

With fast development of mobile phones' hardware and wireless network, above problems may be settled down in near future. We believe using mobile phone as an input system to enrich one's personal digital repository, and providing toolkit for connecting mobile phone with Internet have a large potential industrial market. Our next work will focus on video processing since mobile phone can record a clip of video and send them to server in realtime now. Video-based reconstruction[18] and optimization[19] method will be studied first and results will be used to improve the quality of face texture generation and 3D face model adjustment.

Acknowledgements

This work is co-supported by Key NSF Project on Digital Olympic Museum (Grant no: 60533080), 863 project (grant no: 2006AA01Z303), and Intel/University research grant on 3D HCI and Realistic Avatars Based on Computer Vision. We'd like to give thanks to Dr. You Kin Choong, Sucianto Prasetio, Seah Peck Beng, Clara. Their contributions to this work are invaluable. Dr. Gaoqi He revised my description in detail.

References

[1] Xin, L., Wang, Q., Tao, J.H., Tang, X.O., Tan, T.N., Shum, H.: Automatic 3D Face Modeling fromVideo. In: Proc. of IEEE ICCV, pp. 1193–1199 (2005)
[2] Wen, Z.C., Liu, M., Cohen, J., Li, K.Z., Huang, T.: Low Bit-rate Video Streaming for Face-to-Face Teleconference. In: Proc. of IEEE International Conference on Multimedia and Expo, pp. 309–318 (2004)
[3] Worrall, S.T., Sadka, A.H., Kondoz, A.M.: 3-D facial animation for very low bit-rate mobile video. In: Proc. of 3rd International conference on 3G Mobile Communication Technology, pp. 371–375 (2002)
[4] Ekman, P., Friesen, W.V.: Facial Action Coding System. Consulting. Psychologist Press, Palo Alto, CA (1978)
[5] Eisert, P.: MPEG-4 facial animation in video analysis and synthesis. Journal of Imaging Systems and Technology 2(3), 27–34 (2003)
[6] Tao, H., Chen, H.H., Wu, W., Huang, T.S.: Compression of MPEG-4 facial animation parameters for transmissionof talking heads. IEEE Transactions on Circuits and Systems for Video Technology 9(2), 264–276 (1999)
[7] http://www.arm.com
[8] http://cn.renesas.com
[9] Andrea, L., Ames, D.R., Nadeau, H.L.M.: VRML2.0 Sourcebook. John Wiley and Sons Inc., New York (1997)
[10] http://www.blaxxun.com
[11] http://jcp.org/en/jsr/detail?id=184

[12] Zhang, Z.Y.: A Flexible New Technique for Camera Calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(11), 1330–1334 (2000)
[13] Harris, C., Stephens, M.J.: A combined corner and edge detector. In: Proc. of Alvey Vision Conference, pp. 147–152 (1988)
[14] Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, Cambridge (2003)
[15] Gong, Y., Sakauchi, M.: Detection of regions matching specified chromatic features. Journal of Computer Vision and Image Understanding 61(2), 263–269 (1995)
[16] Pan, Z.G., Zhu, J.J., Hu, W.H.: Interactive Learning of CG in Networked Virtual Environments. Computers and Graphics 2(29), 273–281 (2005)
[17] Jacob, S., Tomas, A.M.: iPACKMAN: High-Quality, Low-Complexity Texture Compression for Mobile Phones. In: Proc. of Graphics Hardware, pp. 63–70 (2005)
[18] Wu, Q., Tang, X.O., Shum, H.: Patch Based Blind Image Super Resolution. In: Proc. of IEEE ICCV, pp. 709–716 (2005)
[19] Shan, Y., Liu, Z., Zhang, Z.Y.: Model-Based Bundle Adjustment with Application to Face Modeling. In: Proc. of IEEE ICCV 2001, pp. 644–651 (2001)

Analysis of Role Behavior in Collaborative Network Learning

Xiaoshuang Xu[1,2], Jun Zhang[2], Egui Zhu[3], Feng Wang[2], Ruiquan Liao[4], and Kebin Huang[2]

[1] School of Computer, Huazhong University of Science and Technology, Wuhan, P.R. China
[2] Dep. of Educational Sci. and Tech., Huanggang Normal University, Huanggang, P.R. China
[3] School of Education, HuBei University, Wuhan, P.R. China
[4] Faculty of Petroleum Engineering, Changjiang University, Jingzhou, P.R. China
{xxsh99,zhangjun,Wangff}@hgnc.net

Abstract. In the collaborative learning environment, role behavior brings on occurrence, development, and disappearance of collaborative learning. In this paper, we first introduce WF-net to describe role behavior in collaborative learning, and then indicate relations among some notions in collaborative learning. We focus on dynamic organization of learning pattern by awareness of role. We represent a mechanism of generating tasks oriented role to ensure persistence of collaborative learning.

Keywords: CSCL; Awareness of role; WF-net; Object.

1 Introduction

Computer supported collaborative learning (CSCL) is changing the traditional ways with the rapid development and effective application. Owing to implementation of CSCL, learners, evaluators and tutors can fulfill their tasks on the basis of a collaborative environment.

In recent years, a number of CSCL applications have appeared. For example, WEB CT at the University of British Columbia has developed a multimedia platform for learners [1]; Virtual-U, a cyberspace campus, can establish a collaborative learning group based on different roles [2]; GRIDCOLE, a Grid Computing Environment, is convenient for learners to access to grid resources, for educators to integrate applications, and for them to participate in collaborative learning applications [3].

The applications above provide multi-mode collaborative mechanism for learners in network environment; however, there still exist several factors to be considered when applied in practice: (1) Learning patterns in the network usually include tutorship, fellowship, and individual and so on. Sometimes learner prefers to finish his task individually in the collaborative environment. Users should be granted privilege to select learning patterns instead of the assigned one during learning. (2) When a user need find a suitable pattern to learn collaboratively, the awareness of role should be

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 562–572, 2008.

provided appropriately. So he can get rid of his isolated feeling and interface with other users in the environment. (3) If learning tasks are generated for a learner to attain his goal, the knowledge background of learners, such as learning level, learning motivation and learning style should be taken into account in the collaborative environment. Learning tasks improper to a learner is likely to result in the failure.

Considering the factors above, in this paper we analyze role behavior in collaborative network learning, indicate any learning organization must have its occurrence, development, and disappearance, and disclose circulation of generating tasks; our contributions can be summarized as follows:

(1) We introduce WF-net to describe role behavior in collaborative learning;

(2) We analyze the organizing of a learning pattern by awareness of role, and make sure aim, condition, and procedure awareness of role;

(3) We provide a mechanism of generating tasks oriented roles, which enable a CSCL system to run continually.

This paper is organized as follows. Section 2 defines notions such as knowledge object, role, and member and so on. Section 3 introduces Petri net and describes of organizing the learning pattern. Section 4 presents how to generate tasks oriented roles in the learning organization. In Section 5 we give a CSCL example implemented on the basis of our thought. Finally, we make conclusions in Section 6.

2 Concepts in the Collaborative Network Learning

In the following parts, we introduce some definitions related to collaborative learning using formal description. To begin,we present definitions of object that will be used in the following discussions.

Definition 1. An object stores its state in fields (variables or data) and exposes its behavior through methods (functions)[4]. Let A be an object, d, one of its variables, denoted as $A.d$; and f, one of its functions, denoted as $A.f$(parameter list),where the parameter list gives necessary data when it is called.

Definition 2. A knowledge tree is a tree where a node is identified for a basal knowledge. Let t be a knowledge tree, we denote $V(t)$ and $E(t)$ the sets of nodes, edges respectively, by $R(t)$ its root node.if $(x, y)\square E(t)$,it means that learning the knowledge node y is the premise condition of learning the knowledge node x. Usually we say $V(t)$ is a knowledge domain[5] w.r.t tree t.

Definition 3. A knowledge unit is a learning object for users. Let u be A knowledge unit, $u.t$,a variable in object u ,denoted that u is connected with knowledge tree t; $u.scope$ is a subset of $V(u.t)$,contains knowledge nodes and represents the learning goal in learning object u; $u.range$, a subset of $u.scope$,means that the learning result. When knowledge unit u is generated, $u.range$ is a null set because nobody learns it. After a learner finish knowledge unit u, $u.range$ represents his knowledge construction. In a knowledge unit other data such as text, graph, audio, and video are contained possibly.

Definition 4. A Tool is an organism of different objects. A Tool is also an object, different tools are provided for users aiming to accomplish the desired learning procedure collaboratively.

Definition 5. A Task is a set of objects correlative to knowledge tree. Particularly, a learning task is a set of knowledge units.

Definition 6. A role is chiefly a semantic functional activity which is constructed according to the specific job [6-7].The roles embody the authority and responsibility and reflect the duty during executing tasks with tools in a system.

Definition 7. An assignation is 3-tuple (*r,tk,tl*),where *r* is a role object, *tk* is a task object, and *tl* is a tool object. Tuple (*tk,tl,r*) means that role *r* is assigned to complete task *tk* with tool *tl*.

Let *r* be a role object, we suppose that tasks *tk* is accompanied with its tool *tl*, we denote *r.execute(tk,tl)={a1,a2,…,an}*,where *a1, a2,… ,an* are different assignations, that the return is a set of assignations. It means after role *r* completes task *tk* with tool *tl*, more tasks with respective tool are assigned to different roles.

In the collaborative environment, there are different roles such as learners, evaluators, tutors and so on, if a role calls function *execute()* by interfacing with a real person, it is called as a real role; otherwise if a role does automatically with nobody, it is called as a virtual role.

If *r* is a real role object, it is enough to constructed human machine interface according to tuple *(tk,tl,r)*.During function *r.execute(tk,tl)* is called, *tk* and *tl* are visualized so that *r* uses tool *tl* to finish task *tk* conveniently. So the collaborative environment oriented to role *r* is established by assignations to it.

Definition 8. An entity is a distinctly identified object. A member is an entity possessing a set of real roles which executes tasks by interfacing with the unique user in a collaborative learning environment; Similarly, An agent is an entity possessing a set of virtual roles which execute tasks automatically without being interfaced with anybody in a collaborative learning environment. If *E001* is an identifier for an entity, either a member or an agent, both role number and role type possessed entity *E001* usually assigned by a proper way. We denote *Role(E001)* the set of role objects possessed by the entity, and *Knowledge(E001)* the set of knowledge nodes mastered by the entity. By the definition, a member can play several roles simultaneously in a real collaborative environment, and it must be mapped to a person, usually called as user. For example, if a member, mapped to a user with identifier *M0002*, but only possesses evaluator object *e* to evaluate the learning task finished by others but also possesses tutor object *t* to assign new learning tasks to them, so *Role(M002) ={e, t}*.On the other hand, For *A003* is an identifier for an agent, *Role(A003)* is the set of possessed roles, and *Knowledge(A003)* is the set of mastered knowledge nodes. In a real collaborative environment, an agent can play one and more virtual roles simultaneously.

3 Organization of Collaborative Learning

To provide a complete presentation of procedure of organizing collaborative learning we review first the basic points from Petri net theory that are used in the later. The

Petri net is a directed bipartite graph with two node types called places and transitions. Directed arcs connected with the nodes. Places are showed by circles and transitions by rectangles. Zero or more tokens are included in the Place, represented by dots in circles. Places in the set correspond to conditions; transitions in the set correspond to activities.

Definition 9. A Petri net structure is triple $N\ (P, T, F)$ [8]:

(1)$P= \{Pl, \ldots Pn\}$ is a finite set of places with $n>0$;

$T= \{tl, \ldots, tm\}$is a finite set of transitions with $m>0$ and $P \cap T= \Phi$.

(2)$F \in (P \times T) \cup (T \times P)$ is the flow relation, a mapping representing arcs between places and transitions. The arcs represented by F prescribe pre- and post- relations for places and transitions.

(3)$p \in P, t \in T, \cdot t=\{p|(p,t) \in F\}$is termed the pre-set of a transition t, and $t\cdot=\{p|(t,p) \in F\}$ is called the post-set of t . The pre-set and post-set for a place p are defined similarly as the sets of transitions incident on s and following s, respectively.

Definition 10. A Petri net $PN\ (P, T, F)$ is a WF-net (workflow-net) if and only if [9]:

(1) PN has two special places: i and o. Place i is a source place and place o is sink place, satisfying $\cdot i= \Phi$, $o\cdot= \Phi$.

(2)A transition t is added to PN which connects place o with i(i.e. $\cdot t=\{o\}$ and $t\cdot=\{i\}$),the resulting Petri net is strongly connected.

3.1 Role and Its WF-Net

By the definition 6 and 7, a role, belonged to either a member or an agent, is capable of performing complex activities. WF-net is identified as a model to describe activities in a system. Therefore, the behaviors of roles can be defined by WF-net. In a real collaborative environment, we find that WF-net for any role has similarly structure, its behaviors are divided into 3 stages: awareness of role, execution of task, and dismission of awareness (See Fig.1). The signification of every transition in Fig.1 is listed in Table 1.

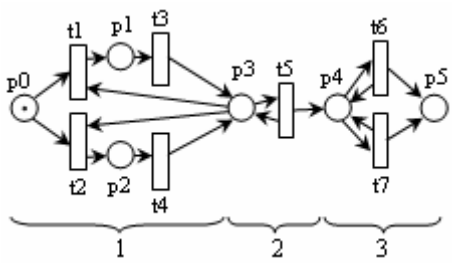


Fig. 1. Behaviors of roles with WF-net

Table 1. The signification of transitions

t	Signification
t_1	Request for awareness
t_2	Respond to awareness
t_3	Decide whether to accept
t_4	Acknowledge decision
t_5	Execute tasks
t_6	Request for dismission
T_7	Respond to dismission

3.2 Aim of Awareness

A role should be entitled to select learning patterns instead of the assigned one during learning. Usually organization of learning pattern is to build relations among roles by

awareness, thus roles can complete tasks cooperatively at the next stage. Relations among roles consist of ones between roles. Before a role possessed by an entity decides to complete some tasks, it should tries to find another role in order that the former assign tasks to by the latter in an environment. If the former fails to do, it should keep waiting. For example, a learner, as a real role, wants to select another role to evaluating its learning result in the network, it requires finding an appropriate evaluator and building mutual relations between them. If any, the learner can begin to learn successfully in the collaborative environment. However, sometimes the learner cannot find any of its proper evaluators, or all evaluators refuse to its request, the learner has to wait until invited by a proper evaluator. Thereof, we declaim that organization of collaborative learning is awareness of role. Awareness of role is corresponding to stage 1 in WF-net (see Fig.1).

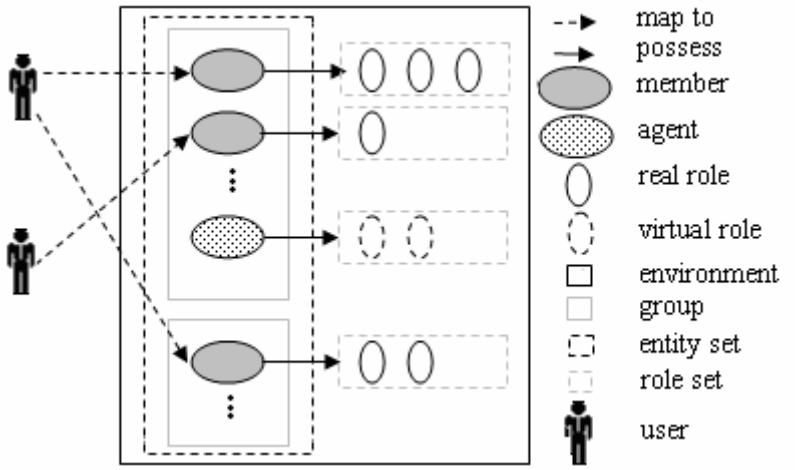


Fig. 2. Relations among definitions in section 2

An environment mainly consists of entity; many definitions in section 2 are correlative in an environment. We present the relations of user, member, agent, entity, role, and group in the collaborative environment in Fig.2. A user may be mapped to different members in several groups at the same time. As a member, he can plays different real roles in one group. When an agent plays different virtual roles in one group, it is not mapped by any user and finishes tasks automatically. Any active role lives in real mode or virtual one, the live mode depends on requirements of the collaborative software and intelligent level of key algorithms. For example, a learner, usually as a real role played by a member in the system because the collaborative software aims at helping students to master knowledge. However, if a learner is a virtual role, student users will not adapt to the collaborative software. Sometimes we create a virtual role, possessed by an agent, to check up a multiple-choice test automatically. However, it is very difficult for a virtual evaluator to give correct score according to an electronic paper from students; in this case, the evaluator must be in real mode and played by a member in an environment.

3.3 Condition of Awareness

During awareness of role as above, it is important for a role to find another. As we all know, an arbitrary role is belonged to an entity. The entity, either a member or an agent, is distinct by its identifier. Let an entity have identifier *E001*, then *Role(E001)* denotes a set of its roles ,and *Knowledge(E001)* denotes a set of its knowledge nodes. So we can know the background information of any role by means of querying properties of the entity which possesses the role. When a role tries to find a proper group to join in, it may check respectively up those roles who respond it according to their background information from corresponding entities, then decide whether to join. So awareness of role means to understand entity in a collaborative group.

We assume that *r* be a role. Before *r* tries to find a proper role, it gives constrained conditions and broadcasts them in the collaborative environment (see transaction t1 in fig.3).Only other roles satisfying conditions can respond to *r* and build relation with *r*. The constrained conditions mainly embody three aspects: role type, knowledge relations and entity set. Let *q* be one of role objects who can respond *r*. So role type indicates *q* must be an object of the role required by *r*; knowledge relation indicates that entity possessing *q* must match to the set of knowledge nodes given by *r*; knowledge relation's value is in set $\{\subset, \subseteq, =, \supset, \supseteq\}$ .entity set indicates that *q* must be longed to one of entities limited by *r*.

For example, Let role *r* be a learner, the entity possessing *r* identified by *E001*, and the entity possessing *q* identified by *E002* .When *r* need awareness of evaluator, *q* has responded to *r*. So *q* must be an evaluator. Meanwhile, if *q* is a qualified evaluator for *r,* knowledge relation $Knowledge(E001) \subseteq Knowledge(E002)$ must be satisfied. Otherwise *q* will fail to evaluate *r* because knowledge of entity *E002* is less than that of entity *E001*. If *r* prefers to evaluators coming from a special part of all entities, *q* must be one element of the limited entity set.

Especially, Let role *r* be a learner, role *q*, *q*' be two evaluators, member *M001* possess both *r,* and member *M002* possess *q*'. When *r* need awareness of evaluator, *q* can respond to *r*. if *q* has become the evaluator for *r*, then member *M001* has to evaluate himself, thus an individual learning has been built in a collaborative environment. On the other hand, if *q*' becomes the evaluator instead of *q*, member *M002* should bear the task of evaluating member *M001* after member *M001*, as a learner, finish a learning task, thus a collaborative learning has been built in a collaborative environment. In a word, there is not a distinct gap between individual learning and collaborative learning; they can transfer each other by means of role behavior in a collaborative environment.

3.4 Procedure of Awareness

We present the procedure by an example. Let role *r* be a learner, role *q* be an evaluators, member *M001* possess *r*, member *M002* possess *q*, and *U* be the set of all entities in an environment. Fig.3 shows corresponding parts of WF-nets of both *r* and *q*. As a sponsor of organization collaborative learning, *r* needs awareness of evaluator. If *q* can respond to *r, q* is one of responders. We show steps of awareness as following:

Step 1: Member *M001* broadcasts its requesting for evaluator role in the group, Role *r* becomes active, and its *t1* is fired (see Fig.3): *r* gives its constrained conditions with three parameters: *T*, *R*, and *S*, where *T*'s value is role type (i.e. evaluator), *R*'s value is one element ($\subseteq$) from set $\{\subset, \subseteq, =, \supset, \supseteq\}$, and *S* dictates the limited entity set satisfying $S \subseteq U$.

Step 2: Member *M002* receives the requesting message, checks every role object in *Role(M002)* and finds *q* is an evaluator without much business. Role *q* becomes active, and its *t2* is fired (see Fig.3): *q* tests whether relation *Knowledge (M001)* $\subseteq$ *Knowledge(M002)* is satisfied, and whether *M002* is in entity set *S*, if any, *q* responses to role *r*.

Step 3: Member *M001* receives the response messages from some roles; it selects *q* from all responders by some algorithm, and build the relation with *q*. Then *r*'s *t3* is fired (see Fig.3): *r* decide to send acceptance message to *q*, and refusal message to the rest. In its next step, one token will appear in *p3*.The token in *p3* means that *r* is ready to complete tasks or to generate complex organization by keeping awareness. On the other hand, Member *M001* doesn't receive any response within limited time, *r*'s *t3* also is fired and *r* restores its WF-net to the original status.

Step 4: Member *M002* receives acceptance message, and build the relation with *r*. Similarly,. *q*'s *t4* is fired (see Fig.3): *q* acknowledges the decision from *r* .In its next step one token will appear in *p3*. The token in *p3* means *q* is ready to cooperate with *r* or to generate complex organization by keeping awareness. For those entities who receive refusal message, transition *t4* of corresponding evaluator role also is fired: the role restores its WF-net to the original status.

A correct schedule of awareness is the dot line showed in Fig.3: the dot line starts from *t1*of *r*, passes by *t2* of *q*, arrives to *t3* of *r*, and end at *t4* of *q*. After awareness between two arbitrary roles, all roles establish collaborative relations. So they may complete tasks regularly in the next stage; the detail will be analyzed in section 4.

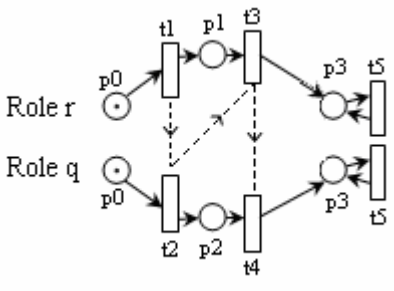


Fig. 3. Awareness of role between *r* and *q*

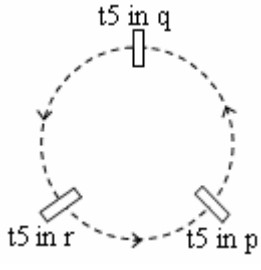


Fig. 4. Collaboration among *q*, *r* and *p*

3.5 Dismission of Awareness

If common goal of roles has been attained, collaboration among roles will terminate. We should unchain relations of keeping awareness between two arbitrary roles, and dismiss the old learning organization, so they can participate in awareness of role repeatedly, a new learning organization may be generated.

Dismission of awareness is corresponding to stage 3 in the WF-net (see Fig.1).

Let *r*, *q* be two roles of keeping awareness each other, one of them (e.g. *r*), fires its t5 as a sponsor in order to dismiss their awareness, then *q* responds to *r* and fires its t6. Thus the procedure of dismissing between them has finished (see Fig.3). All roles dismiss awareness, the learning organization vanishes completely.

4 Collaboration in the Learning Environment

We think that any complex collaboration consists of collaboration between roles, so the collaborative learning environment is based on collaboration between roles. Roles start cooperating from their initial assignations.

4.1 Collaboration between Roles

Collaboration happens in stage 2 of role's WF-nets. Awareness as above has built mutual relations between roles, so every role knows how to distribute new tasks to different roles after executing its task.

Let role *r* be a role, tasks *tk* be assigned to *r* with its tool *tl*, and the entity possessing *r* be identified by *E001*. Suppose that function *r.execute*() is called, we have

$$r.execute(tk,tl)=\{(r1,tk1,tl1),\ (r2,tk2,tl2),\ (r3,tk3,tl3)\}\ . \tag{1}$$

where *r1*, *r2* and *r3* are three roles, *tk1*, *tk2* and *tk3* are their tasks respectively; *tl1*, *tl2* and *tl3* are their corresponding tools. The result, a set of assignations, means that three new tasks and their corresponding tools are generated, and distributed to *r1*, *r2* and *r3*.This function indicates that collaboration among three roles consists of one between r and any of *r1*, *r2*, and *r3* in the learning environment.

During the collaboration, the kind of task is various. For a learning role, the tasks can range from clicking buttons, typing a text, and designing 3D work to having an exam online and so on; for an evaluator, the tasks include giving score according to learning result, and adding new knowledge nodes mastered by a learner; for a tutor, he often decides the next learning according to a learner's status and a knowledge tree.

In the collaborative environment, there are different roles such as learner, evaluator, tutor, administrator, monitor, recorder and so on. The complexity of collaboration lies on but only the kind, number and awareness of role, but also the kind, number and difficulty of task during the collaborative learning.

4.2 Collaboration between Roles

If assignations of tasks among roles recycle, a collaborative learning keeps continuing. For example, let *r* be a learner, *p* be an evaluator, and *q* be a tutor. Three entities possessing *r*, *p* and *q* are identified by *E001*, *E002* and *E003* respectively. The knowledge background of three entities is based on knowledge tree *t*. Fig.4 shows transaction *t5* in WF-net of each role.

In stage 1 of their WF-nets, three roles finish awareness. Role *q* builds relations with *r* such that *r* can receive new learning task from *q*; *r* builds relations with *p* such

that p can evaluate the learning result of r; p builds relations with q such that q can make learning plan according to evaluation and knowledge tree t.

In stage 2 of their WF-nets, by the awareness result of role we have

$$\begin{cases} q.execute\,(tk_i^q, tl_i^q) = \{(r, tk_i^r, tl_i^r)\} \\ r.execute\,(tk_i^r, tl_i^r) = \{(p, tk_i^p, tl_i^p)\} \\ p.execute\,(tk_i^p, tl_i^p) = \{(q, tk_{i+1}^q, tl_{i+1}^q)\} \end{cases} . \tag{2}$$

where $i=1, 2\ldots n$.

In formula (2), symbols tk_i^q and tl_i^q denote that task tk_i^q with tool tl_i^q is assigned to q in the ith time, rest of symbols have analogous meaning. At the beginning of collaboration (i.e. $i=1$), both tk_1^q and tl_1^q are default parameters given by the system.

Obviously, assignations of tasks among roles creates a cycle (See fig.4).After q finishes its tutoring task, it generates a new learning task for r. After learning generates, a new evaluating task is sent to p. After examination of learning, p gives a new tutoring plan to q. After analyze the tutoring plan according to background of the learner and the knowledge tree, q lists another new learning task for r. So a collaborative learning can keep continuing.

When function $p.execute(tk_i^p, tl_i^p)$ is called, a learning task tk_i^r, a set of knowledge units, is finished by learner r and sent from learner r to evaluator p as an element of set tk_i^p .For any knowledge unit $u \in tk_i^r$, $u.range$ is the learning result given by evaluator p, is denoted as knowledge nodes mastered by learner r in knowledge unit u, it is added to *knowledge(E001)* during evaluating. In a form, we have

$$knowledge(E001) \leftarrow (\bigcup_{u \in tk_i^r} u.range) \cup knowledge(E001) \ . \tag{3}$$

Thus the updated *knowledge* (*E001*) is match up to current knowledge level of role r in order that tutor q lists a new learning task for r according background of the learner and goal of the learning.

4.3 Termination of Collaboration

Let r be a role, task tk be assigned to r with its tool tl, and the entity possessing r be identified by *E001*. Suppose that function *r.execute*() is called, we have

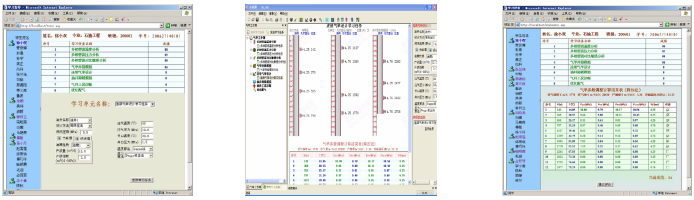
$$r.execute\ (tk,tl)=\emptyset\ . \tag{4}$$

Formula (3) means to stop generation of any new task, and terminate collaboration between roles. In the example as above, if a tutor finds that his learner has mastered all knowledge nodes (i.e. $V(t)=$ *knowledge* (*E001*)), he should stop teaching the learner, and dismiss the awareness; the learner does nothing because he can not get any task from his tutor, but respond to dismissing from his tutor. All roles stop

generating task, the collaborative learning terminates completely; all roles dismiss awareness, the learning organization vanishes completely.

5 An Example

We have developed GLLS (Gas-Lift Learning System) to help students major in petroleum engineering. The gas-lift is a basic mode of exploiting petroleum, so it is important for students to design and analyze of gas-lift wells. They should master knowledge nodes such as valves placement, gas allocation, simulation of pressure and temperature in wells, valves diagnosis, gas-lift unloading, optimization of gas allocation in gas-lift block and so on. In the system, we introduce three roles: learner, evaluator and tutor. Any user can play all roles. Firstly, those users build learning organization by awareness of role, Secondly, every tutor creates learning task and send to his student (See fig.5 (a)); the learner receives task and begin to finish it t(See fig.5(b));the evaluator give scores to learners (See fig.5(c)).Since the given scores are background knowledge for a learner, the tutor creates another new learning task according them. Finally, users can dismiss the organization randomly. The learning organizations are different owing to the fact that members possess roles.



(a) (b) (c)

Fig. 5. Three roles are Learning Collaboratively in GLLS

6 Conclusions and Future Work

The goal of the research presented in this paper is to disclose how roles organize a learning pattern in the collaborative environment. We point out relations among notions during collaborative learning. We introduce WF-net to describe role activity, and find a learning pattern is established after awareness of role. We provide a method of generating tasks oriented roles, which enable a CSCL system to run continually in a collaborative learning pattern.

In the future work, we are interested in the learning organization of adaptively, changeability and complexity. One of the challenges in further researches is to analyze competition in collaborative learning. Both collaboration and competition should appear in learning applications. We will try to discuss the cause and mechanism of generating competition among roles, and focus on both relations and differences with collaboration in collaborative network learning.

References

1. Bruce, M., Chan, L.K.S.: Co-operative learning in integrated classroom. Curriculum and Teaching 6(1), 48–52 (1991)
2. Celine, B.: Complementarity of information and quality of relationship in cooperative learning. Social Psychology of Education 4(3/4), 335–357 (2001)
3. Bote-Lorenzo, M.L., Vaquero-Gonzalez, L.M., Vega-Gorgojo, G.: GRIDCOLE: a Grid Collaborative Learning Environment. In: 2004 IEEE International Symposium on Cluster Computing and the Grid, pp. 105–112. IEEE Press, Chicago (2004)
4. The Java Tutorials, http://java.sun.com/docs/books/tutorial/java/concepts/object.html
5. Guo, H., Sun, J.-m.: Research and design of intelligent teaching model and collaborative learning mechanism. In: The 7th International Conference on Computer Supported Cooperative Work in Design, pp. 465–469. NRC Research Press, Ottawa (2002)
6. Mattas, A.K., Mavridis, I.K., Pangalos, G.I.: Towards Dynamically Administered:Role-Based Access Control. In: Mařík, V., Štěpánková, O., Retschitzegger, W. (eds.) DEXA 2003. LNCS, vol. 2736, pp. 494–498. Springer, Heidelberg (2003)
7. Joshi, J.B.D., Bertino, E., Latif, U., et al.: A Generalized Temporal Role-based Access Control Model. IEEE Trans. on Knowledge and Data Engineering 17(1), 4–23 (2005)
8. Reising, W.: Petri-nets: An introduction. Springer, New York (1982)
9. van der Aalst, W.M.P.: The application of Petri nets to workflow management. The Journal of Circuits, System and Computers 8(1), 21–66 (1998)
10. Eberspacher, H., Joab, M.: A role-based approach to group support in a collaborative learning environment. In: Proc. of the Fifth IEEE International Conference on Advanced Learning Technologies, pp. 64–65. IEEE Press, Taiwan (2005)
11. Zhang, R., Guo, L.-j., Liu, Z.: A novel a administrative model for collaborative learning system. In: Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, pp. 4159–4162. IEEE Press, Hong Kong (2007)

Survey on Real-Time Crowds Simulation

Mohamed 'Adi Bin Mohamed Azahar[1], Mohd Shahrizal Sunar[2], Daut Daman[2], and Abdullah Bade[2]

Faculty of Computer Science and Information System
Universiti Teknologi Malaysia
81310 Skudai, Johor, Malaysia
[1]vahnzenqi@gmail.com, [2]{shahrizal,daut,abade}@utm.my

Abstract. The simulation of human massive crowds play an important role in real-time application such as games and walkthrough system. This kind of applications can provide an immersive feeling of life into the static scene and enhance the reality of the system. In recent years, there are many significant research and techniques have been developed, mainly focus on the entertainment industry for both real-time and non-real time rendering. This paper will give an overview on crowd behavior in real-time crowd simulation. The work will also cover numerous crowd modeling and rendering techniques.

Keywords: Crowd Simulation, Crowd Rendering, Computer Graphics.

1 Introduction

The wide use of computer graphics in education, entertainment, games, simulation, and virtual heritage applications has led it to become such an important area of research. In simulation, it is important to create an interactive, complex, and realistic virtual world so that the user can have an immersive experience during navigation through the world [1]. As the size and complexity of the environments in the virtual world increase, it becomes more necessary to populate them with peoples, and this is the reason why rendering crowds in real-time is crucial.

In general, crowd simulation constitute of three important areas. There are realism of behavioral [2], high-quality visualization [3], and convergence of both areas. Realism of behavioral is mainly targeted for simple 2D visualizations because most of the attentions are concentrated on simulating the behavior of the group. . High-quality visualization is regularly used for movie productions and computer games. In this area, behavior is not really critical. The most important thing is how we can produce very convincing visual. Convergences of both areas are mainly used for application like training systems. In order to make the training system more effective, the element of valid replication of the behavior and high-quality visualization will be added.

As we make comparison with non real-time system, developer for real-time crowd simulation requires to consider various kind of challenges. One of the challenges is to provide efficient management approach at every level of simulation in order to ensure, the agents composing a crowd should look different, move different, act different, and so forth, similar to the real world. Another challenge is the increasing of

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 573–580, 2008.

demand on computational resources because the systems need to compute behavior, take and process input not known in advance, and to render large and varied crowds instantaneously.

The purpose of this paper is to present a review on crowd behavior, previous researches on crowd simulation, and crowd modeling and rendering technique. Section two will present about real-time crowd simulation timeline. Section three will discuss about crowd behavior in details as well as crowd modeling and rendering techniques. Afterward, we conclude our paper with some future research directions pertaining the area.

2 Real-Time Crowd Simulation Timeline

Real-time crowd simulation is a process of simulating the movement of a large number of animated characters or agents in the real-time virtual environment. Crowd movement in certain cases requires the agents to coordinate among themselves, follow after one another, walking in line or dispersing using different directions. All of these actions will contribute to the final collective behavior of the crowds that must be achieved in real-time. Unlike non-real-time simulations which is able to know the full run of the simulated scenario, real-time simulations have to react to the situation as it unfolds in the moment. Real-time rendering of a large number of 3D characters is also a challenge, because it can exhaust the system resources quickly even for a powerful system [4].

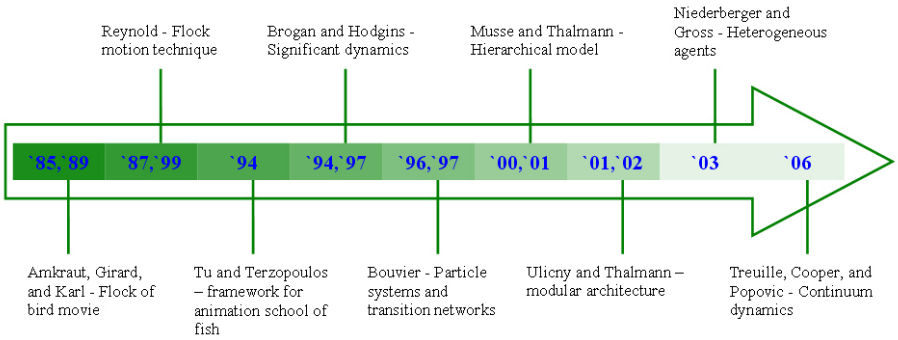


Fig. 1. This figure shows timeline on previous work for crowd simulation. More detail on development of crowd simulations are discussed below.

The first procedural animation of flocks of virtual birds was shown in the movie by Amkraut, Girard, and Karl called Eurhythmy, for which the first concept [5] was presented at The Electronic Theater at SIGGRAPH in 1985 and the final version was presented at Ars Electronica in 1989. The flock motion was achieved by a global vector force field guiding a flow of flocks. A behavioral animation of human crowds is based on foundations of group simulations of much more simple entities, notably flocks of birds [6] and schools of fish [7].

In his pioneering work, Reynolds [6] described a distributed behavioral model for simulating aggregate motion of a flock of birds. The technical paper was accompanied by an animated short movie called "Stanley and Stella in: Breaking the Ice" shown at the Electronic Theater at SIGGRAPH '87. The revolutionary idea was that a complex behavior of a group of actors can be obtained by simple local rules for members of the group instead of some enforced global condition. The flock is simulated as a complex particle system, with the simulated birds, called boids, being the particles. Each boid is implemented as an independent agent that navigates according to its local perception of the environment, the laws of simulated physics, and the set of behaviors. The boids try to avoid collisions with one another and with other objects in their environment, match velocities with nearby flock mates, and move toward a center of the flock. The aggregate motion of the simulated flock is the result of the interaction of these relatively simple behaviors of the individual simulated birds. Reynolds later extended his work by including various steering behaviors as goal seeking, obstacle avoidance, path following, or fleeing [8], and introduced a simple finite-state machines behavior controller and spatial queries optimizations for real-time interaction with groups of characters [9].

Tu and Terzopoulos proposed a framework for animation of artificial fishes [7]. Besides complex individual behaviors based on perception of the environment, virtual fishes have been exhibiting unscripted collective motions as schooling and predator evading behaviors analogous to flocking of boids.

Brogan and Hodgins [10] and [11] simulated group behaviors for systems with significant dynamics. Compared to boids, a more realistic motion is achieved by taking into account physical properties of motion, such as momentum or balance. Their algorithm for controlling the movements of creatures proceeds in two steps: first, a perception model determines the creatures and obstacles visible to each individual, and then a placement algorithm determines the desired position for each individual given the locations and velocities of perceived creatures and obstacles. Simulated systems included groups of one-legged robots, bicycle riders, and point-mass systems.

An approach similar to boids was used by Bouvier et al. [12] and [13] to simulate human crowds. They used a combination of particle systems and transition networks to model crowds for the visualization of urban spaces. At the lower level, attractive and repulsive forces, analogous to physical electric ones, enable people to move around the environment. Goals generate attractive forces, obstacles generate repulsive force fields. Higher level behavior is modeled by transition networks with transitions depending on time, visiting of certain points, changes of local population densities, and global events.

Musse and Thalmann [14] and [15] presented a hierarchical model for real-time simulation of virtual human crowds. Their model is based on groups, instead of individuals: groups' are more intelligent structures, where individuals follow the groups specification. Groups can be controlled with different levels of autonomy: guided crowds follow orders (as go to a certain place or play a particular animation) given by the user in run-time; programmed crowds follow a scripted behavior; and autonomous crowds use events and reactions to create more complex behaviors. The environment comprises a set of interest points, which signify goals and way points; and a set of action points, which are goals that have some actions associated. Agents move between way points following Bezier curves.

Recently, Ulicny and Thalmann [16] and [17] presented a crowd behavior simulation with a modular architecture for multiagent system allowing autonomous and scripted behavior of agents supporting variety. In their system, the behavior is computed in layers, where decisions are made by behavioral rules and execution is handled by hierarchical finite-state machines.

Another work was exploring group modeling based on hierarchies. Niederberger and Gross [18] proposed an architecture of hierarchical and heterogeneous agents for real-time applications. Behaviors are defined through specialization of existing behavior types and weighted multiple inheritances for creation of new types. Groups are defined through recursive and modulo based patterns. The behavior engine allows for the specification of a maximal amount of time per run in order to guarantee a minimal and constant frame rate.

Most recently, a real-time crowd model based on continuum dynamics has been proposed by [19]. In their model, a dynamic potential field integrates global navigation with moving obstacles, efficiently solving for the motion of large crowds without the need for explicit collision avoidance.

3 Crowds Behavior

Crowd behavior modeling plays a major role in the domain of safety science and architecture. Application such as crowd evacuation simulators need to apply crowd behavior modeling to show believable visual and accurate result. Their objective is to assist designers to understand the relation between the surroundings space and human behavior [20].Many areas in computer graphics that can utilize crowds’ behavior, such as:

1. Training and simulator system
2. Game industry
3. Simulation and animation

Crowd behaviors modeling play an essential part in police and military simulator systems used for training. Simulator such as *CACTUS* developed to help in planning and training for public order incidents such as large demonstrations. Game industry nowadays still did not fully integrating virtual crowds in their game play environment because of the need for high computational resource and costly production process. Recently, there are some game genres that starting to change this situation, such as real-time strategy genre that integrating groups of armies which will give direct effect on the game play. In order to create an immersive simulation application using crowds in virtual environments, researchers need to deal with various aspect of crowd simulation such as behavioral animation, environment modeling, and crowd rendering [4]. Nowadays, many movies and animation start to use crowd simulation in their production. One of the most advanced crowd animation system for non-real-time productions is Massive, this system was used to create battle scenes for The Lord of the Rings movie trilogy.

There are two major crowd behavior models that have been used successfully, which are PetroSim behavioral model [21] and physically based behavioral model [22]. PetroSim is a system for simulation of evacuations of outdoor urban areas in real

time. Physically based behavioral model was proposed by Braun et al. [22] for simulation of crowd evacuations from internal environments including several rooms and obstacles.

4 Modeling and Rendering

The tricky part when dealing with thousands of characters is the quantity of information that needs to be processed. Simple approaches, where virtual humans are processed one after another without specific order will produce high computational cost for both the CPU and GPU. This is the reason why data that flows through the same path need to be grouped for an efficient use of the available computing power. This paper will present simple architecture that able to handle and sort virtual human related data into grouped slots. For the best simulation result, characters capable of facial and hand animation are simulated in the area near to the camera to improve believability, while for farther area, less expensive representations are used. Concerning efficiency of storage and data management, database must be used to store all the virtual human-related data. The presented architecture is versatile enough to be used in different scenarios, such as in confined environments and in large-scale environments.

4.1 Modeling

To visualize crowds of virtual humans, thousands of very detailed meshes can be used, such as capable of hand and facial animation. The concept of levels of detail (LOD) is exploited to meet real-time constraints. Levels of detail for virtual humans composing crowd that depend on the location of the camera; a character is rendered with a specific representation, resulting lower rendering cost and adequate visual quality. A type of human such as a woman, man, or child is described as a human template. Each rendered virtual human is derived from a human template, recognized as an instance of a human template. Instances of the same human template must look different using several appearance sets.

In order to perform level of detail, human template must be modeled for each level of detail, these level are deformable mesh, rigid mesh and impostor. A deformable mesh is a representation of a human template composed of triangles. It is enveloping a skeleton of 78 joints, used for animation: when the skeleton moves, the vertices of the mesh follow smoothly its joint movements, similarly to our skin. Unfortunately, the cost of using deformable meshes as the sole representation of virtual humans in a crowd is too excessive. Therefore it is used in limited number and only at the forefront of the camera.

A rigid mesh is a precomputed geometric posture of a deformable mesh, thus sharing the very same appearance. A rigid animation sequence is always inspired from an original skeletal animation, and from an external point of view, both look alike. The gain in speed brought by this new representation is considerable. It is possible to display about 10 times more rigid meshes than deformable meshes. However, the rigid meshes need to be displayed farther from the camera than deformable meshes,

because they allow for neither procedural animations nor blending, and they also have no composited, facial, or hand animation.

Impostors are the less detailed representation, and extensively exploited in the domain of crowd rendering [23] and [3]. An impostor represents a virtual human with only two textured triangles, forming a quad, which is enough to give the wanted illusion at long range from the camera. Impostor is only a 2D image of the posture that is kept for each keyframe, instead of the whole geometry. Since impostors are only 2D quads, normals and texture coordinates from several points of view need to be store, so that, at runtime, when the camera moves, we can display the correct keyframe from the correct camera view point. Since impostor is only 2D representation of the human template, it's usually placed farthest away from the camera.

4.2 Rendering

There are several technique used to speed up rendering process in crowd simulation. Billboarded impostors are one of the methods used to speed up crowd rendering. Impostors are partially transparent textured polygons that contain a snapshot of a full 3D character and are always facing the camera. Aubel et al. [24] introduced dynamically generated impostors to render animated virtual humans. A different possibility for a fast crowd display is to use point-based rendering techniques. Wand and Strasser [25] presented a multiresolution rendering approach which unifies image based and polygonal rendering. An approach that has been getting new life is that of geometry baking. By taking snapshots of vertex positions and normal, complete mesh descriptions are stored for each frame of animation as in the work of Ulicny et al. [26]. A more recent approach to crowd rendering using geometry is through dynamic meshes as presented in the work of de Heras et al. [27], where dynamic meshes use systems of caches to reuse skeletal updates which are typically costly.

Fig. 2. This figure depicts the different stages: at each frame, data are flowing sequentially through each one of them, beginning from the left, to the right [4]

5 Future Work

The Ancient Malacca Virtual Walkthrough [28] is a project that focuses on the modeling and visualization of Malacca city in 15th century. It is based on local and foreign sources, such as the Sejarah Melayu and the descriptions by Portuguese writer, Tome Pires. The focus area of visualization is Central Business and Administrative District of the Malacca Empire. The project is visualized in real-time rendering mode using SGI Onyx 3800 with 16 CPU, 32GB RAM and three Infinite Reality3 graphics pipes. At the moment there are no crowd simulation have been develop for this project. For future work, we will add crowd simulation into this walkthrough system with average

computational load that can be supported by normal personal computer. The challenge of this project is to bring crowd simulation into the walkthrough system on a normal personal computer with the same or better quality compared to super computer. As conclusion, we hope that this research is useful for the crowd simulation system developer and also can give benefit to the computer graphics community.

Acknowledgement

We would like express our appreciation to Malaysian Ministry of Science, Technology and Innovation under eScienceFund grant (01-01-06-SF0066) for financial support of this research. We also like to thanks Creative Application Development Centre (CADC), Multimedia Development Corporation, Cyberjaya, Malaysia for allowing us to use the 3D model taken from Ancient Malacca project and permit us to explore the potential improvement of the project.

References

1. Tecchia, F., Loscos, C., Chrysanthou, Y.: Visualizing Crowds in Real-Time. Computer Graphics Forums (2002)
2. Thompson, P., Marchant, E.: Testing and application of the computer model 'simulex'. Fire Safety Journal 24(2), 149–166 (1995)
3. Dobbyn, S., Hamill, J., O'Conor, K., O'Sullivan, C.: Geopostors: A real-time geometry/impostor crowd rendering system. In: SI3D 2005: Proceedings of the 2005 Symposium on Interactive 3D Graphics and Games, pp. 95–102. ACM Press, New York (2005)
4. Thalmann, D., Musse, S.R.: Crowd Simulation. Springer, London (2007)
5. Amkraut, S., Girard, M., Karl, G.: Motion studies for a work in progress entitled "Eurnythmy". SIGGRAPH Video Review 21 (second item, time code 3:58 to 7:35) (1985)
6. Reynolds, C.W.: Flocks, herds, and schools: A distributed behavioral model. In: Computer Graphics (ACM SIGGRAPH 1987 Conference Proceedings), Anaheim, CA, USA, vol. 21, pp. 25–34. ACM, New York (1987)
7. Tu, X., Terzopoulos, D.: Artificial fishes: Physics, locomotion, perception, behavior. In: Computer Graphics (ACM SIGGRAPH 1994 Conference Proceedings), Orlando, FL, USA, vol. 28, pp. 43–50. ACM, New York (1994)
8. Reynolds, C.W.: Steering behaviors for autonomous characters. In: Proceedings of Game Developers Conference 1999, pp. 763–782 (1999)
9. Reynolds, C.W.: Interaction with groups of autonomous characters. In: Proc. Game Developers Conference 2000, pp. 449–460 (2000)
10. Hodgins, J., Brogan, D.: Robot herds: Group behaviors for systems with significant dynamics. In: Proc. Artificial Life IV, pp. 319–324 (1994)
11. Brogan, D., Hodgins, J.: Group behaviors for systems with significant dynamics. Autonomous Robots 4, 137–153 (1997)
12. Bouvier, E., Guilloteau, P.: Crowd simulation in immersive space management. In: Proc. Eurographics Workshop on Virtual Environments and Scientific Visualization 1996, pp. 104–110. Springer, Heidelberg (1996)
13. Bouvier, E., Cohen, E., Najman, L.: From crowd simulation to airbag deployment: Particle systems, a new paradigm of simulation. Journal of Electrical Imaging 6(1), 94–107 (1997)

14. Musse, S.R.: Human Crowd Modelling with Various Levels of Behaviour Control. PhD thesis, EPFL, Lausanne (2000)
15. Musse, S.R., Thalmann, D.: A hierarchical model for real time simulation of virtual human crowds. IEEE Transactions on Visualization and Computer Graphics 7(2), 152–164 (2001)
16. Ulicny, B., Thalmann, D.: Crowd simulation for interactive virtual environments and VR training systems. In: Proc. Eurographics Workshop on Animation and Simulation, pp. 163–170. Springer, Heidelberg (2001)
17. Ulicny, B., Thalmann, D.: Towards interactive real-time crowd behavior simulation. Computer Graphics Forum 21(4), 767–775 (2002)
18. Niederberger, C., Gross, M.: Hierarchical and heterogeneous reactive agents for real-time applications. Computer Graphics Forum 22(3) (Proc. Eurographics 2003) (2003)
19. Treuille, A., Cooper, S., Popovic, Z.: Continuum crowds. ACM Transactions on Graphics 25(3), 1160–1168 (2006)
20. Okazaki, S., Matsushita, S.: A study of simulation model for pedestrian movement with evacuation and queuing. In: Proc. International Conference on Engineering for Crowd Safety 1993 (1993)
21. Barros, L.M., da Silva, A.T., Musse, S.R.: Petrosim: An architecture to manage virtual crowds in panic situations. In: Proceedings of Computer Animation and Social Agents, pp. 111–120. ACM Press, New York (2004)
22. Braun, A., Bodman, B.J., Musse, S.R.: Simulating virtual crowds in emergency situations. In: Proceedings of ACM Symposium on Virtual Reality Software and Technology — VRST 2005, Monterey, CA, USA, pp. 244–252. ACM, New York (2005)
23. Tecchia, F., Loscos, C., Chrysanthou, Y.: Visualizing crowds in real-time. Computer Graphics Forum 21(4), 753–765 (2002)
24. Aubel, A., Boulic, R., Thalmann, D.: Real-time display of virtual humans: Levels of detail and impostors. IEEE Transactions on Circuits and Systems for Video Technology 10(2), 207–217 (2000)
25. Wand, M., Strasser, W.: Multi-resolution rendering of complex animated scenes. Computer Graphics Forum 21(3) (Proc. Eurographics 2002) (2002)
26. Ulicny, B., de Heras Ciechomski, P., Thalmann, D.: Crowdbrush: Interactive authoring of real-time crowd scenes. In: Proc. ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA 2004), pp. 243–252 (2004)
27. De Heras, P., Schertenleib, S., Maïm, J., Thalmann, D.: Realtime shader rendering for crowds in virtual heritage. In: Proc. 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST 2005) (2005)
28. Sunar, M.S., Mohd Zin, A., Tengku Sembok, T.M.: Effective Range Detection Approach for Ancient Malacca Virtual Walkthrough. The International Journal of Virtual Reality 5(4), 31–38 (2006)

TS-Animation: A Track-Based Sketching Animation System

Guangyu Wu, Danli Wang, and Guozhong Dai

Institute of Software, The Chinese Academy of Sciences, Beijing 100190, China
guangyuwu81@hotmail.com, danli@iso.com,
Guozhong@admin.iscas.ac.cn

Abstract. Computer has been widely used in classroom. Yet most of them can only provide static slides in the project mode, which are not attractive to children, especially to the younger ones. On contrast animations are not only attractive but also expressive. Using some animations in classroom makes the children more active and enhances their learning efficiency. However, most of current animation tools are either too complicated to common users, or specific domain based. Aimed at removing the complexity barrier and allowing nonprofessional users to create a wide variety of animations quickly, we propose a sketching animation system. A kind of track-based method is adopted to make creating animations as easy as drawing. Motion coordination is simplified by providing the tool of motion time warping. The interface of our system allows most of the important motions to be set by pen gestures.

Keywords: Animation, Sketching, coordination, track-based.

1 Introduction

Since (at least) 1970s, computers entered classrooms as teaching assistants. Computer-Aided Teaching has been of great benefits to teachers and students. The electronic lecture systems, such as Microsoft's PowerPoint, Apple's Keynote and so on, are used widely, which is often considered as the biggest technology revolution in classrooms [1]. However, both of them don't support freehand sketching, which is a familiar, efficient, and natural way of expressing certain kinds of ideas, particularly in the classroom. Teachers get used to sketching on the blackboard with chalk. Yet this archetypal behavior is largely unsupported by current software used in classroom. More and more researchers have recognized that such slides make the students keep in passive receiving, not active thinking, because the students could not get any response from their teachers and are not involved in the lectures, since teachers can not interact with their students in time [2]. Differently, pen-based interaction has some advantages of writing easily, drawing freely, taking up just one hand, and so on. Hence, pen-based interaction is particularly suitable for lecturing in class, since there is less text input, more sketches and more operations in class, and since the teacher do not have much time to sit on a chair in class in order to attract his students.

Although sketching rough diagrams is a useful tool in a classroom, sometimes static diagrams are insufficient .Even very a simple animation can be more expressive, it is a

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 581–592, 2008.

convenient way to express moving visual images, it can represent dynamic concepts, and it can make information more attractive and engaging[3], especially to the children. Animation can also make students move active in the learning process. Research on Pedagogy indicates that active learning could remarkably enhance students' learning efficiency. And abundant interactions can give students participant feelings and, hence, benefit students' active learning and promote teaching effects.

The traditional way to create animation is to draw a series of images in a flip-book. This is not only time consuming but also tedious. Although computer has been used to help animators remove this tedium, one problem remains: animation tools and skills are still in the hands of a small number of animation makers. Current animation tools have extremely complex interfaces with many, specialized methods for generating motion from static specifications [4]. This may be easier than designing every frame by hand, but it is still too complicated for most of the teachers.

Aimed at these issues, this paper proposed a track-based sketching animation system, which implements several methods to improve access to animation through a simple interface for animation sketches. By focusing on motion-by-example and gesture input, we hope to make animating to be an easy work, so that common users can create animations without any complicated barriers.

This paper begins by reviewing the wealth of related work in animation tools. Then we describe the track-based approach of animation creating, and motion coordination. In Section 4 we give an outline of our system, its interface, a simple using scenario and an informal evaluation. Conclusion and future works are given in Section 5.

2 Related Work

Researchers have long tried to create informal Animation sketching tools. The first informal animation sketching tool was Baecker's Genesys system [5]. In recent years, the use on free–form strokes for creating 2D free–form objects and animation has become popular. Richard C. Davis and James A. Landay conducted field studies that investigated potential users' needs in an "informal animation sketching" tool .And they implemented a system that supports the most important needs, but they neglected the coordination of motions, which is also an important aspect to create an expressive animation. Tomer Moscovich and John F. Hughes discussed several methods to aid sketching of animation, and had implemented a system as a test-bed for these methods [6]. They proposed the approach of motion-by-example and even synchronization, but their solution for object coordination is still incomplete. Their work is the most similar to ours. There are also several complicated systems such as TicTacToon [7], Fabian Di Fiore and Frank Van Reeth's Multi–Level Sketching Tool [8]. Although these systems can be used to create complicated and vivid animations by professional users, they are too complicated for common users. For education, there have been noteworthy systems that generate domain-specific 2D animations from sketches of mechanical systems [9] and math equations [10].We focus instead on general-purpose animations.

Free–form strokes are also being used for rapidly designing3D free–form objects and 3D animations. Igarashi, Matsuoka, & Tanaka present a gesture–based sketching interface (TEDDY) to quickly and easily design 3D free–form models out of 2D

sketched strokes [11]. Later, they introduce spatial key-framing, a technique for performance-driven character animation [12]. This technique is especially useful for imaginary characters other than human figures because it does not rely on motion-capture data. On the contrary Matthew Thorne, David Burke, & Michiel van de Panne presented a novel system for sketching the motion of a character [13]. Although sketch has been used in 3D widely, creating 3D objects is still more difficult than 2D, and some of them can only support very limited motions.

We attempt to create a simple sketching animation tool allowing inexperienced users to create a wide variety animation quickly. And we believe 2D animation to be simpler and more accessible to the common users.

3 Track-Based Sketching Animation

Based on the work of Tomer Moscovich and John F. Hughes [6], we proposed an approach of track-based sketching animation. Our approach to create an animation consists of three phases drawing objects, setting motions, and coordinating motions. In the first phase users draw some objects, which either can be a simple stroke or a composition of stokes (see Figure 1). In the setting animations phase, our basic approach is motion-by-example [6]. All basic animations (Translation of objects, Rotation of objects, Scale of objects) are represented by a track. The user simply moves the object as desired, and the timing and position information are recorded as a track. The motion can be played back following the recorded track producing a simple animation. In the coordinating motions phrase, the main task is to make animations coordinate correctly by adjusting their time.

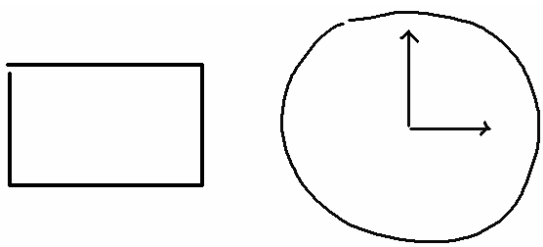


Fig. 1. The left is a rectangle containing one stroke only. The right is a clock composed by 5 strokes.

3.1 Object Representing

We represent a complicated object with a tree structure, of which every leaf-node is a single stroke. All 5 leaf-nodes of the clock in Figure1 are shown in Figure 2. When select operation is done, the selected strokes constitute a new node, if more than one stroke is selected. Figure 3 shows the growing process, of the tree for the clock. The growing process is transparent to the user, who just selects some stokes and sets motions to them without knowing the tree structure.

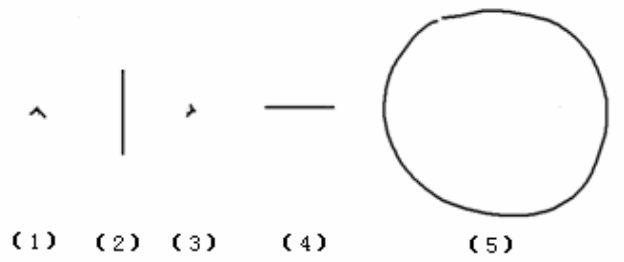


Fig. 2. The leaf-nodes of the Clock in Fig. 1

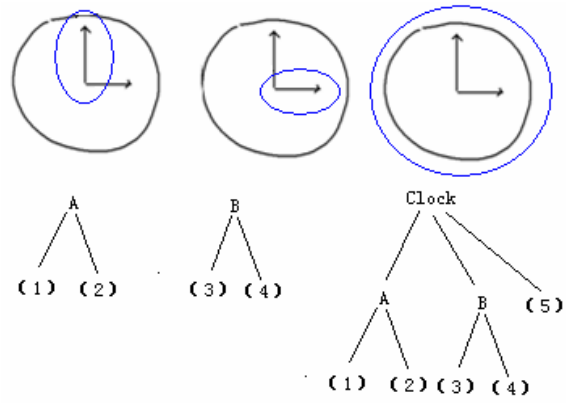


Fig. 3. The use's selecting operations and the tree nodes created. The leaf-nodes are shown in Figure 2.

3.2 Setting Motions

We adopt the approach of track-based animation. In track-based animation, the user simply grabs an object of interest, and moves it about as he/she likes. The position and timing information is recorded. Every motion is recorded as a track. In Figure 4, the user sets two motions to a rectangle stroke, i.e., a rotation and a translation. The motions can be then played back, following the track of the animator's hand. This is an easy way to create motion [6].

The user draws a stroke S_0 on the canvas, and then sets N motions for it. We use T_i to represent the i-th motion（1≤i≤N）. Then we can calculate the transform matrix M for S_0 at time t.

$$\mathrm{M(t)} = \left(\prod_{i=1}^{n} M_i(t)\right) \cdot M_p(t). \tag{1}$$

$M_i(t)$ is the transform matrix of T_i at time t, and $M_p(t)$ is the transform matrix of the parent-node of S_0 at time t. A track (Figure 5) for a motion is recorded as

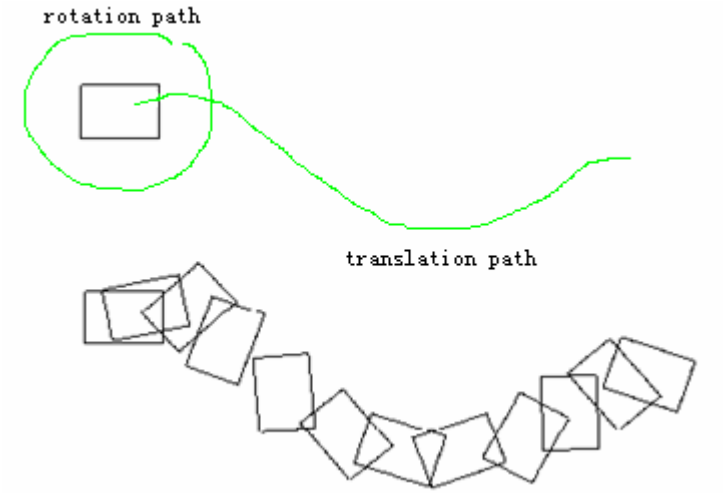


Fig. 4. The user drags a rotation path and a translation path to a rectangle

Path= $P_1\ P_2 \ldots P_n$, where $P_i = (x_i, y_i, t_i)$. Then the transform matrix of T_i at time t can be computed using P_j and P_1, where j=max(i , where 1≤i≤n and $t_i \leq$t) . We get the shape of S_0 at time t using Formula 2.

$$S(t) = S_0 \cdot M(t) . \tag{2}$$

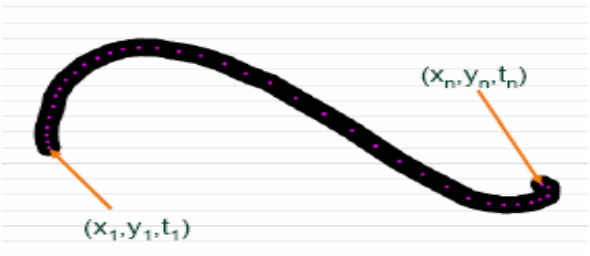


Fig. 5. A path with time information

3.3 Motion Coordination

One problem the user runs into when using track-based animation is that of coordination. In frame-based techniques, the animator can cause things to occur simultaneously, or in some other temporal relationship such as cause-and-effect simply by drawing them in the appropriate frame [6]. In track-based animation, the animator must rely on his sense of timing and response speed. For example, in Figure 4 the user attempts to make the rectangle begin to rotate and translate at the same time, and also

wants them to stop at the same time. First the user sets the rotate motion, then he/she attempts to add the translate motion. After he/she sees the rectangle beginning to rotate, he/she drags it to move, until he/she sees it stop rotating he/she stops moving it. The progress depends on the user's sense, and inevitably the translate motion starts and stops behind the rotate motion. In order to help the user to succeed in this task, we provide a time edit tool, which allows the user to reset the start and end time of a motion. Assume the origin time of the points on a motion path is ($t_1\ t_2 \ldots t_m$), t_1 is the start time and t_m is the end time. After the user resets time, the start time and the end time become to be t_1^* and t_m^* respectively. We use the Formula (3) to reset all points' time on that motion path.

$$t_i^* = t_1^* + (t_m^* - t_1^*)\frac{(t_i - t_1)}{(t_m - t_1)} \quad (3)$$

Sometimes, resetting the start and end time is not enough. As shown in Figure 6, the user wants the rectangle to move along the sine curve, while the arrow to rotate around the origin point. After reset their start and end time, they don't collaborate correctly in the mid time although they start and stop at the same time.

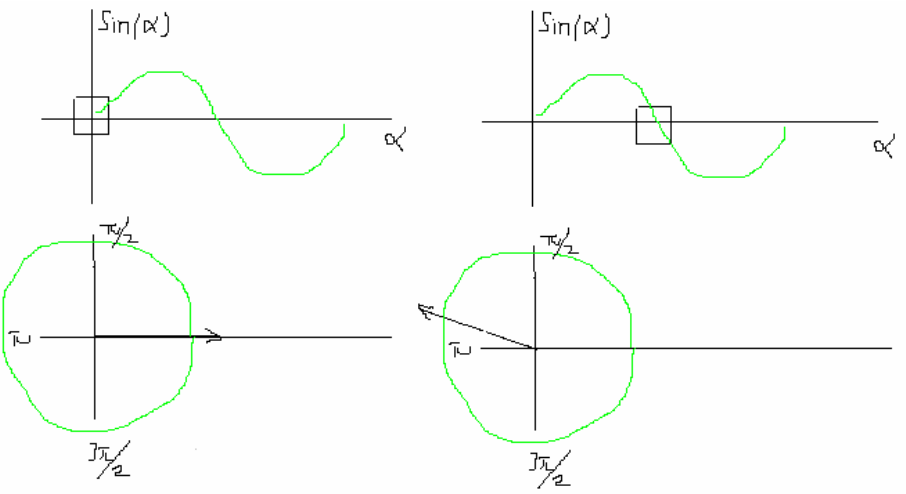


Fig. 6. Left: the rectangle's translation motion and the arrow's rotate motion. Right: the rectangle moves to the position of $\alpha = \pi$, but the arrow does not rotate to the degree of π.

We provide another tool to solve this problem, key-points. The user can set several key points for a motion, seeing Figure 7. There are three key points on the translation motion, and three key points on the rotate motion. Then the user adjusts the time of the key points. He/she just need to adjust the collaborative key points to the same time, then the motions can collaborate very well. There is no limit for number of key

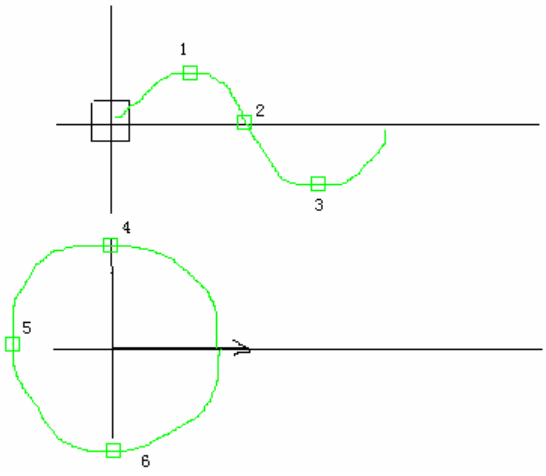


Fig. 7. The key points set respectively for the translation motion and the rotate motion. Key-point 1 and Key-point 4 have the same time. Key-point 2 and Key-point 5 have the same time. Key-point 3 and Key-point 6 have the same time.

points the users can set for a motion. More key points a motion have, more well the motion collaborates with others. However for most of simple motions, 3 to 5 key points are enough.

4 System Description

We implement a system supporting track-based sketching animation, in which animation can be created using any of a variety of devices that provide the experience of freehand drawing while capturing pen movement. We represent a stroke with the set of points produced by the drawing implement between the time it contacts the surface (mouse-down) and the time it breaks contact (mouse-up). A stroke can be either a part of an object or a path of an animation. The animation methods supported by our system includes translation of objects, rotation of objects, scale of objects, appearance/disappearance, repeating motion and the composition of them. These include most of the important methods used in an informal animation discussed by Davis RC and Landay JA [4].

4.1 Gesture and Sketch Process

In order to make the system easy to use, we designed several gestures, selecting object, setting key-point, deleting motions etc. When a stroke is inputted, the function of Gesture and Sketch process is to identify the stroke as a gesture, a drawing or a motion (Figure 8).

The first part of Gesture and Sketch Process is a preprocessor, which gets the user's operation information. Pen is the main input device of this system. Hence, no matter what mode the system is in, drawing or animation setting, the pen' s movements and

some other attributes are logged. The useful information for the following task includes all the points on the pen's trace, which are organized as a stroke, the speed, the pen's angle of inclination and so on. The strokes are real-time collected by the preprocessor and sent to the classifier in time. Besides, it also does some work to make the strokes visually more appealing without changing their basic shape. We refer the method described by Tevfik Metin Sezgin [14].

The following part is a classifier, which classify the strokes into 3 categories: gestures, drawings and motions, according to the context information and a group of pre-defined rules. The last part is composed of three processors which are in charge of motions, drawings and gestures respectively.

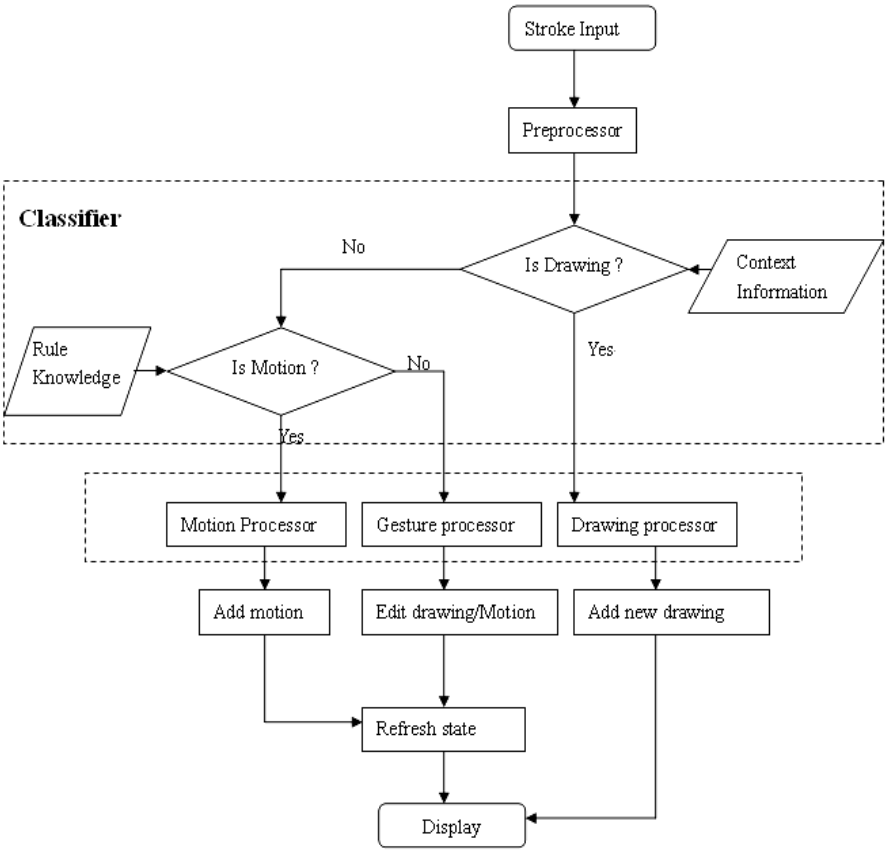


Fig. 8. Gesture and Stroke Process

4.2 User Interface

We have designed a simple animation user interface that will allow users to draw objects and create motions that our system supports. Setting repeat times and editing time of motions are also easy.

There are three areas, a blank canvas, a motion selecting area and a motion time editing area (Figure 9). Users draw objects and set motions for them on the blank canvas. To define motions, the user presses "Start", and all drawings and modifications are recorded. In order to avoid confusion, the paths of motions are not visible, unless it is selected to edit. When an object is selected, the selection widget in Figure 10 appears and all the motions of the object are shown in the motion selecting area. In the motion selecting area**,** user can select a motion just by double-click on it, delete a motion by a simple delete-gesture, and reset its repeat num just by resetting the num on its top-left corner. When a motion is selected, its start time, end time and key points' time appear in the time editing area. The user drags them to reset their time. The time editing area can show two motions, so the user can make them coordinate easily.

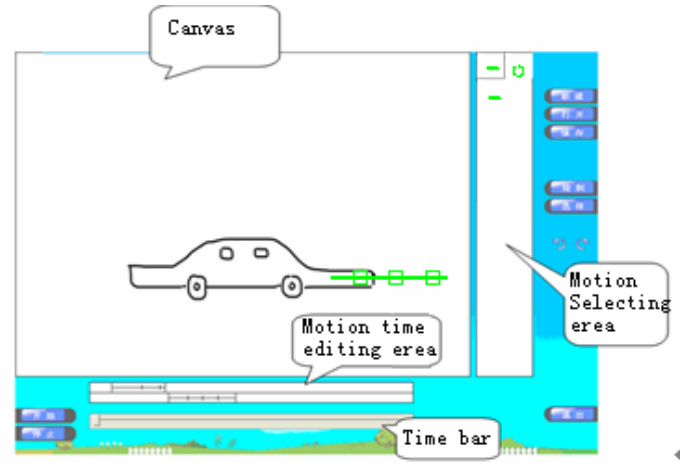


Fig. 9. The animation interface

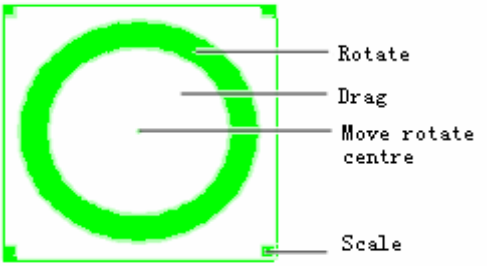


Fig. 10. The selection widget with multiple control zones. Users may specify a variety of motions (such as translating or rotating) or other operations (such as moving the center of rotation). This widget is similar to that of KSketch.

4.3 A Simple Using Scenario

A teacher had prepared a lecture for the relationship knowledge among the sun, the earth and the moon. The moon encircles around the sun while earth encircles around

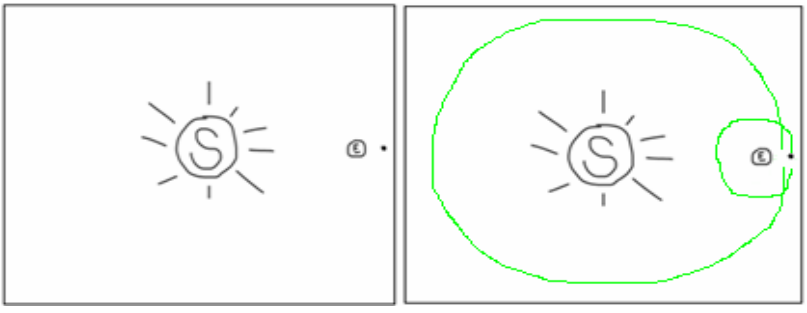


Fig. 11. Left: The teacher draws the sun, the earth and the moon. Right: The track for the moon and the track for the earth.

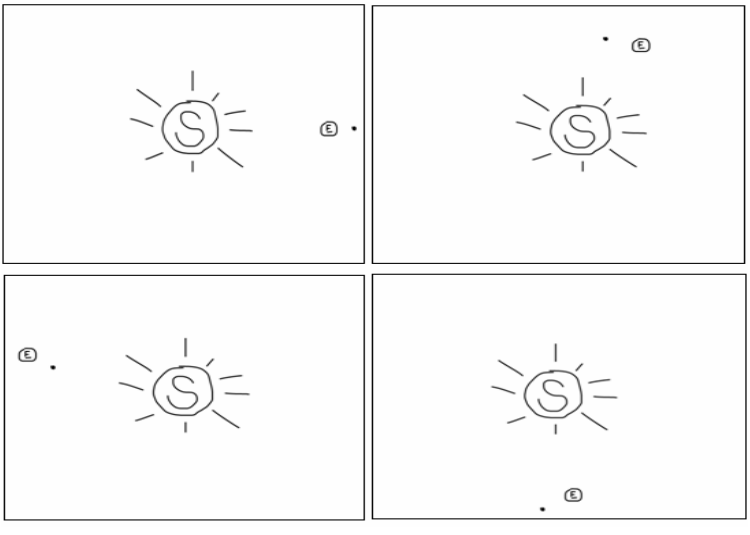


Fig. 12. The playing of the animation

the sun. In order to give a vivid image of this relationship, the teacher decided to use an animation. First he draw the three celestial bodies on the blank canvas (figure 11 left), then he set an encircling path for the moon around the earth and an encircling path for the earth around the sun (Figure 11 right). At last he reset their motion time to make them synchronize well. It's done.

In the classroom, the teacher showed the animation just by pressing the "start" button. The animation is shown in Figure 12. He also can pause or restart the animation at anytime. Through this vivid animation, the students get to understand the relationship more deeply.

4.4 Informal Evaluation

We have invited several users to use and evaluate our system, including two courseware makers who have some animation experience, and three common users.

After learning and practicing for a short time, they can all get hold of it. They find it interesting, easy to use and can preserve a lot of time for simple animations. The courseware makers said they are willing to use it as an aid. Besides, they give us a lot of valuable suggestions to improve it. (1)One courseware maker said it is too simple to support more complicated and attractive animations. (2)Composition of different motions relies on their setting order, which may confuse the user. (3) Morphing motion is important to create vivid animations. (4) Interactive animation is a complement. In the future work we will take these issues into account.

5 Conclusion and Future Work

In this paper, we propose a track-based sketching animation system. It enables teachers to create simple animations for their class in order to promote active learning. The track-based approach removes the complexity barrier while supporting most of the needs of informal animation. And the system provides a simple interface which enables the user to do any freehand drawing as he wants and set motions easily. With this system a teacher can create animation just with a pen very easy, even children can create their own animations.

However, the work is far from complete. There are still many works we need do in the future. First morphing and Free-form Skeleton are also important ways for animation making, so next we will investigate ways to extend this system to support morphing and Free-form Skeleton while keeping the interface simple. Second, it is difficult to make complicated coordination with several objects. Adjusting the time of too many key points is tedious. Third the integration of sketching animation with existing method may be a way to make the animation more expressive and vivid. For example, one could use track-based animation to control the movement of characters, while using key-frame techniques or morphing for controlling their shapes.

Acknowledgement

The research is supported by National Grand Fundamental Research 973 Program (Grant No. 2002CB312103), the National Natural Science Foundation of China (Grant No. 60373056), National Facilities and Information Infrastructure Foundation for Science and Technology(Grant No.2005DKA46300-05-32) and the CAS Pilot Project of the National Knowledge Innovation Program (Grant No. KGCX2—YW—606).

References

1. Bonwell, Charles C., James A.: Eison, Active Learning: Creating Excitement in the Classroom. ASHE-ERIC Higher Education Report. Washington, D.C. (1991)
2. Bransford, John, Brown, Ann, Rodney, C. (eds.): How People Learn: Brain, Mind,Experience, and School. Committee on Developments in the Science of Learning,Commission on Behavioral and Social Sciences and Education, National Research Council. National Academy Press, Washington (1999)

3. Park, O., Hopkins, R.: Instructional Conditions for Using Dynamic Visual Displays: A Review. Instructional Science 21, 427–448 (1993)
4. Davis, R.C., Landay, J.A.: Informal animation sketching: Requirements and design. In: Proc. of the AAAI 2004 Fall Symp. on Making Pen-Based Interaction Intelligent and Natural, Washington (2004)
5. Baecker, R.: Picture-Driven Animation. In: Proceedings of the AFIPS Spring Joint Computer Conference, vol. 34, pp. 273–288 (1969)
6. Moscovich, T., Hughes, J.F.: Animation sketching: An approach to accessible animation. Technical report, Brown University ((2003)
7. Fekete, J.-D., Bizouarn, É., Cournarie, E., Galas, T., Taillefer, F.: TicTacToon: a paperless system for professional 2D animation. In: Proceedings of SIGGRAPH, pp. 79–90 (1995)
8. Di Fiore, F., Van Reeth, F.: A Multi–Level Sketching Tool for "Pencil–and–Paper" Animation. In: Sketch Understanding: Papers from the 2002 American Association for Artificial Intelligence (AAAI Spring Symposium), Palo Alto (USA), March 25-27, 2002, pp. 32–36 (2002)
9. Davis, R.: Sketch Understanding in Design:Overview of Work at the MIT AI Lab. In: AAAI Spring Symposium on Sketch Understanding, pp. 24–31 (2002)
10. LaViola, J.J., Zeleznik, R.C.: MathPad2: A system for the Creation and Exploration of Mathematical Sketches. Proceedings of SIGGRAPH 23(3), 432–440 (2004)
11. Igarashi, T., Matsuoka, S., Tanaka, H.: Teddy: a sketching interface for 3D freeform design. Computer Graphics, 33(Annual Conference Series), 409–416 (1999)
12. Igarashi 1,3, T., Moscovich 2, T., Hughes, J.F.: 2Spatial Keyframing for Performance-driven Animation Eurographics/ACMSIGGRAPH Symposium on Computer Animation (2005)
13. Thorne, M., Burke, D., Van De Panne, M.: Motion Doodles: An Interface for Sketching Character Motion. In: Proceedings of SIGGRAPH (2004)
14. Sezgin, T.M., Stahovich, T., Davis, R.: Sketch Based Interfaces: Early Processing for Sketch Understanding. In: Proc. 2001 Workshop Perceptive User interface (PUI 2001), pp. 1–8. ACM Press, New York (2001)

Dynamic Axial Curve –Pair Based Deformation

M.L. Chan and K.C. Hui

Department of Mechanical and Automation Engineering
The Chinese University of Hong Kong
Shatin, Hong Kong

Abstract. Deformation of 3D objects plays an important role in computer graphics, simulation and computer-aided design. Using a deformation tool, a simple geometric model can be deformed to take useful and intuitive shapes. The axial deformation technique allows a 3D object to be deformed by adjusting the shape of an axial curve. However, due to lack of control on the local coordinate frame, unexpected twist may result. The axial curve-pair based deformation technique provides a scheme for controlling the local coordinate frame intuitively. Nevertheless, achieving a physically viable deformation relies very much on the experience and skill of the user in manipulating the shape of the curve-pair. The dynamic axial curve-pair based deformation technique enhances the system by incorporating a special mass-spring model for the 1-dimensional curve-pair structure. Movement of the point masses of the mass spring system deforms the embedded curve-pair, which in turn deforms the associated geometric shape. The proposed technique is particularly useful for the design and animation of soft objects such as animals and characters.

Keywords: Axial Curve-pair deformation, dynamic deformation.

1 Introduction

Deformation of 3D objects plays an important role in computer graphics, simulation and computer-aided design. Using a deformation tool, a simple geometric model can be deformed to take meaningful and creative shapes. Freeform deformation (FFD) [1 - 4] deforms an object by embedding the object in a space which can be warped, and thereby deforming the embedded shape. Similar to FFD, other deformation tools [5, 6] also require manipulating a large number of parameters to achieve a desired deformation. The axial deformation technique was introduced by Lazarus et al. [7] which deforms a 3D object by adjusting the shape of an axial curve. However, the lack of control on the local coordinate frame of the axial curve may result in unexpected twist of the object.

Hui [8] proposed a free-form design method using axial curve-pairs. A curve-pair composing of a primary and an orientation curve provides explicit control on the local coordinate frame of the axial curve. By associating 3D objects to the curve-pair, these objects can be stretched, bended and twisted through manipulating the curve-pair. The axial curve-pair deformation technique is an effective tool for manipulating complex shapes in industrial and aesthetic design. However, to deform objects in a physically

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 593–601, 2008.

viable manner, the control points of the axial curve-pair have to be adjusted such that the axial curve-pair is deformed according to physical laws. Incorporating physical properties in the axial curve-pair thus provides a convenient tool for deforming and animating soft objects. The D-NURBS approach [10] effectively models an elastic NURBS curve, but may not be applicable for modeling elastic curve-pair.

In this paper, a dynamic axial curve-pair is adopted. The deformation technique makes use of an energy model, the mass-spring system, to emulate the physical properties of an elastic curve-pair as described below.

2 Physics Based Axial Curve-Pair

In the proposed approach, a mass spring system is used to simulate the physical properties of a curve-pair. This can be achieved by approximating the curve-pair with a mass-spring system. By simply implementing tensile interactions between the neighboring point masses in a curve-pair mass spring structure, the mechanical behaviors of the curve-pair are usually limited. For example, the elastic twist and bending of the curve-pair cannot be modeled [11 - 15].

A special mass-spring model is deigned for the 1-dimensional curve-pair representation. Each point mass is linked to its neighbors by different types of springs in order to simulate different elastic behaviors of the curve-pair. The springs tend to keep the point masses at their initial resting positions. The dynamic behavior of the system is simulated by updating the mass spring system at discrete time-steps. At each step, the spring forces are calculated and applied to the point masses, which respond by accelerating in the direction of the net force based on Newton's second law. The forces exerted on each point mass depend on the current state of the system, which is defined by the location of the point mass, orientation of the local coordinate frames, and external interactions. These forces include internal force due to elasticity, viscosity, gravity, and other external forces and constraints.

Using a forth order Runge-Kutta method with adaptive step-size, the equations of motion are integrated and the position and velocity of each point mass can be obtained [19 - 21]. By interpolating a curve-pair through the mass points of the mass-spring system, a deformed curve-pair is obtained which in turn deforms the objects associated with the curve-pair.

2.1 Framing the Curve-Pair

The curve-pair technique [8] provides explicit control on the local coordinate frame of the axial curve. By associating an object to the curve-pair, the object can be stretched, bended and twisted intuitively through adjusting the control points of the curve-pair.

Denote the primary curve as $\mathbf{c}(t)$, and the orientation curve as $\mathbf{c_D}(t)$, where $t'_{start} \leq t' \leq t'_{end}$, the curve-pair can be constructed as a set of k th-degree B-spline curves: $\mathbf{c}(t) = \sum_{i=0}^{n} \mathrm{N}_{i,p}(t)\, \mathbf{u}_i$ and $\mathbf{c_D}(t) = \sum_{i=0}^{n} \mathrm{N}_{i,p}(t)\, \mathbf{v}_i$, where $\mathbf{u}_i$ is the i-th control point of the primary curve, $\mathbf{v}_i$ is the i-th control point of the orientation curve, $\mathbf{v}_i = \mathbf{u}_i + d$ where d is the offset distance, $\mathrm{N}_{i,p}(t)$ is the B-spline basis function of degree p. When a vertex of a model $\mathbf{s}$ is attached to the curve $\mathbf{c}(t)$, a projection point $\mathbf{s_p} = \mathbf{c}(t)$ is obtained. A local coordinate frame is constructed at $\mathbf{s_p}$ as discussed below.

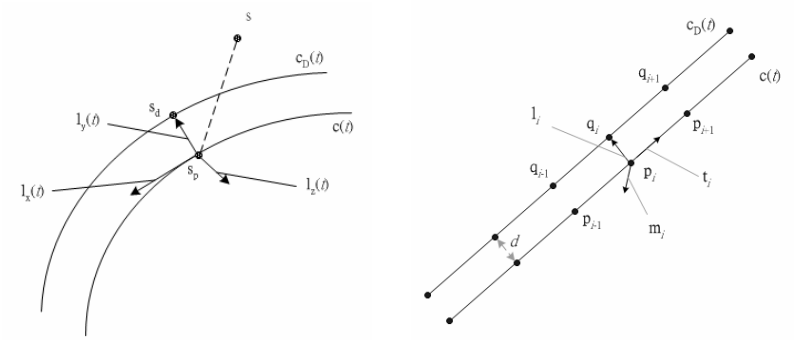


Fig. 1. The local coordinate frame of an axial curve-pair (left) The axial curve-pair model (right)

Using the same parametric value t at $\mathbf{s_p}$, a point $\mathbf{s_d}$ is obtained by projecting the point $\mathbf{c_D}(t)$ onto the plane through the point $\mathbf{s_p}$ with unit normal $\mathbf{c'}(t) / |\mathbf{c'}(t)|$. Hence, $(\mathbf{s_p} - \mathbf{s_d}) \cdot \mathbf{c'}(t) = 0$ and the local coordinate frame is given by

$$\begin{aligned} \mathbf{l_z}(t) &= \mathbf{c'}(t) / |\mathbf{c'}(t)| \\ \mathbf{l_x}(t) &= \mathbf{n}(t) \text{ x } \mathbf{l_z}(t) \\ \mathbf{l_y}(t) &= \mathbf{l_z}(t) \text{ x } \mathbf{l_x}(t) \end{aligned} \tag{1}$$

where $\mathbf{n}(t) = \mathbf{c_D}(t) - \mathbf{c}(t) / |\mathbf{c_D}(t) - \mathbf{c}(t)|$

2.2 Framing the Curve-Pair

Assume the primary curve $\mathbf{c}(t)$ passes through point masses $\mathbf{p}_i$ and the orientation curve $\mathbf{c_D}$(t) passes through point masses $\mathbf{q}_i$. These mass points also satisfy the condition that $\mathbf{q}_i = \mathbf{p}_i + d$, where d is the offset distance.

Given an axial curve-pair ($\mathbf{c}$, $\mathbf{c_D}$), the local coordinate frame at $\mathbf{p_i}$ is given by the unit vectors $\mathbf{l}_i$, $\mathbf{m}_i$, $\mathbf{t}_i$, where $\mathbf{t}_i$ is the polygon tangent [8] at $\mathbf{p}_i$, and $\mathbf{l}_i = (\mathbf{q}_i - \mathbf{p}_i)$ x $\mathbf{t}_i$ $|\mathbf{q}_i - \mathbf{p}_i|$, $\mathbf{m}_i = \mathbf{t}_i$ x $\mathbf{l}_i$.

There are five degrees of freedom for the point $\mathbf{p}_i$ including translation in the directions $\mathbf{t}_i$, $\mathbf{l}_i$, $\mathbf{m}_i$ and rotation about the axis $\mathbf{l}_i$ and $\mathbf{m}_i$. Similarly, there are six degrees of freedom for the point $\mathbf{q_i}$ including translation in the directions $\mathbf{t}_i$, $\mathbf{l}_i$, $\mathbf{m}_i$, and rotations about the axis $\mathbf{t}_i$, $\mathbf{l}_i$ and $\mathbf{m}_i$.

The motion of $\mathbf{p}_i$ is governed by the mass-spring system simulating the bending and stretching of the curve-pair. However, the motion of $\mathbf{q}_i$ is not fully controlled by the system. When pi is modified, $\mathbf{q}_i$ has to be modified correspondingly. Only the rotation axis $\mathbf{t}_i$ of $\mathbf{q}_i$ is governed by the dynamic system to represent the elastic twisting properties of the curve-pair.

Geometric parameters such as the rest length between two control points, the initial angle between two local coordinate frames, and the tangent at a point mass determines the geometry of the mass spring model. Physical parameters including forces, masses, stiffness and stress governs the physical behavior of the system, and which can be obtained from the geometric parameters.

2.3 The Curve-Pair Mass-Spring Model

A mass density m is associated with the point masses $\mathbf{p}_i$ and $\mathbf{q}_i$. The linear stress at a point mass is induced by the variation in linear displacement from the rest-length of a spring along the direction of the spring. The torsional stress at a point mass is induced by the variation in angular displacement from the rest-angle between the lattice edge and lattice frame (Figure 2). To express these stress behavior, spring elements are connected between the point masses. Different spring elements are adopted to represent the different physical behavior of a curve-pair model.

Two types of springs are adopted. They are the metric springs and torsional springs. The structure of the mass-spring model is specially designed for the 1-dimensional curve-pair model. To maintain the linear structure and to capture the tensile stress behavior, metric spring elements are connected between neighboring point masses (e.g. $\mathbf{p}_i$, $\mathbf{p}_{i+1}$) of the mass spring model.

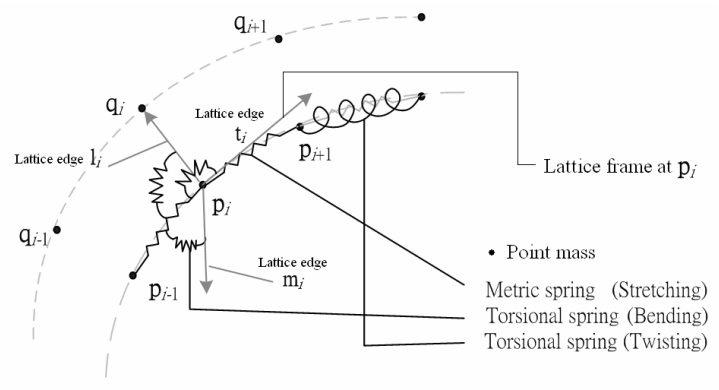


Fig. 2. Structure of mass spring model

Elastic bending is modeled by two types of spring elements. The angular spring elements defined at the angle of the lattice edge and the neighboring local coordinate frame axes provide a controllable bending deformation. To capture the torsional stress behavior, torsional spring elements are connected between the local coordinate frames ($\mathbf{l}_i$, $\mathbf{m}_i$, $\mathbf{t}_i$) at $\mathbf{p}_i$ and ($\mathbf{l}_{i+1}$, $\mathbf{m}_{i+1}$, $\mathbf{t}_{i+1}$) at $\mathbf{p}_{i+1}$.

2.4 Energy Model

Lagrangian formulation is used to formulate the dynamics of the mass-spring system. The Lagrangian equation of motion is expressed as:

$$m\mathbf{r}'' + d\mathbf{r}' + k\mathbf{r} = \mathbf{f}_{external} \tag{2}$$

for a point mass $\mathbf{r}$, where m, d, k are respectively, the mass, damping and the linear stiffness for the point mass $\mathbf{r}$, and $\mathbf{f}_{external}$ is the external force acting on the point mass $\mathbf{r}$.

The third term $k\mathbf{r}$ represents the internal force at the point mass $\mathbf{r}$, and is expressed as $-[\,\mathbf{f}^{\delta}(\mathbf{r})+\mathbf{f}^{\theta}(\mathbf{r})+\mathbf{f}^{\lambda}(\mathbf{r})+\mathbf{f}^{\beta}(\mathbf{r})\,]$ where $\mathbf{f}^{\delta}$, $\mathbf{f}^{\theta}$, $\mathbf{f}^{\lambda}$, $\mathbf{f}^{\beta}$ are torques with rotation angles as showed

in Figure 3. Equation (2) can be written in matrix form to obtain Equation (3), in which **R** represents the 2*N* x 3 dimensional position vector, that is, $[\ \mathbf{p}_0\ \ \mathbf{p}_1 \ldots \mathbf{p}_{N-1}\ \ \mathbf{q}_0\ \ \mathbf{q}_1 \ldots \mathbf{q}_{N-1}]^T$. *N* is half the total number of point masses.

$$M\mathbf{R}'' + D\mathbf{R}' + K\mathbf{R} = \mathbf{F}_{external} \tag{3}$$

In Equation 3, *M* and *D* are 2*N* x 2*N* diagonal mass and damping matrices respectively, *K* is the 2*N* x 2*N* symmetric sparse matrix, and Fexternal is the 2*N* x 3 force vector. Assuming there is no damping, Equation (2) can be simplified to:

$$M\mathbf{R}'' + K\mathbf{R} = \mathbf{F}_{external} \tag{4}$$

The total force on the model is a vector sum of the internal and external forces. The internal forces can be the tensile stresses or torsional stresses. User-interactions, reaction forces and gravity are treated as independent external force in the dynamics formulations. In the next section, the formulations of different forces are described.

2.4.1 Tensile Stress

The tensile stress at a point mass is induced by the difference in the linear displacement along the metric springs connecting neighboring control points. The tensile stress at a point mass **p** with a spring of stiffness K^t and initial length δ connecting its *j*-th neighbor is defined as:

$$\mathbf{F}^{\delta}_i = \mathrm{K}^t_i\ [\ (\ \mathbf{p} - \mathbf{p}_j\ ) + \delta_j\ \mathbf{u}_j\ ] \tag{5}$$

where u_j is the unit vector which is equal to $(\ \mathbf{p} - \mathbf{p}_j\ ) / |\ \mathbf{p} - \mathbf{p}_j\ |$.

2.4.2 Tortional Stress

The torsional stress at a point mass is induced by the difference in angular displacement about the angular springs connecting neighboring local coordinate frames. There are four types of angular deformations:

(i) Bending at $\mathbf{p}_i$ along the angles between $\mathbf{P}_{ij}$ and $\mathbf{t}_i$
(ii) Bending at $\mathbf{p}_i$ along the angles between $\mathbf{P}_{ij}$ and $\mathbf{l}_i$
(iii) Bending at $\mathbf{p}_i$ along the angles between $\mathbf{P}_{ij}$ and $\mathbf{m}_i$
(iv) Twisting about $\mathbf{P}_{ij}$ along the angles between $\mathbf{l}_i$ and $\mathbf{l}_j$

where $\mathbf{P}_{ij}$ is the direction vector $\mathbf{p}_i - \mathbf{p}_j$, and $\mathbf{p}_j$ is the *j*-th neighbor of $\mathbf{p}_i$. For the first three types of angular deformations, the torque at the point mass pi is proportional to the angular displacement $\Delta\theta$, $\Delta\lambda$ and $\Delta\beta$. The torque can be expressed as:

$$\begin{aligned} \tau^{\theta}_i &= K^{\theta}\ (\Delta\theta\ )\ d_{i,j}\ \mathbf{n}_{\theta,i} \\ \tau^{\lambda}_i &= K^{\lambda}\ (\Delta\lambda\ )\ d_{i,j}\ \mathbf{n}_{\lambda,i} \\ \tau^{\beta}_i &= K^{\beta}\ (\Delta\beta\ )\ d_{i,j}\ \mathbf{n}_{\beta,i} \end{aligned} \tag{6}$$

where K^{θ} ,K^{λ} ,K^{β} are the respective angular stiffness, $d_{i,j}$ is the length of vector $\mathbf{P}_{ij}$ and **n** is the normal unit vector which is given by:

$$\begin{aligned} \mathbf{n}_{\theta,i} &= \mathbf{P}_{ij} * \ \mathbf{l}_i / d_{i,j} \\ \mathbf{n}_{\lambda,i} &= \mathbf{P}_{ij} * \ \mathbf{m}_i / d_{i,j} \\ \mathbf{n}_{\beta,i} &= \mathbf{P}_{ij} * \ \mathbf{t}_i / d_{i,j} \end{aligned} \tag{7}$$

where $\mathbf{t}_i$ is the polygon tangent at $\mathbf{p}_i$, and $\mathbf{l}_i = (\mathbf{q}_i - \mathbf{p}_i) \times \mathbf{t}_i \mid \mathbf{q}_i - \mathbf{p}_i \mid$, $\mathbf{m}_i = \mathbf{t}_i \times \mathbf{l}_i$. The torques are then converted to virtual forces on the neighboring point mass.

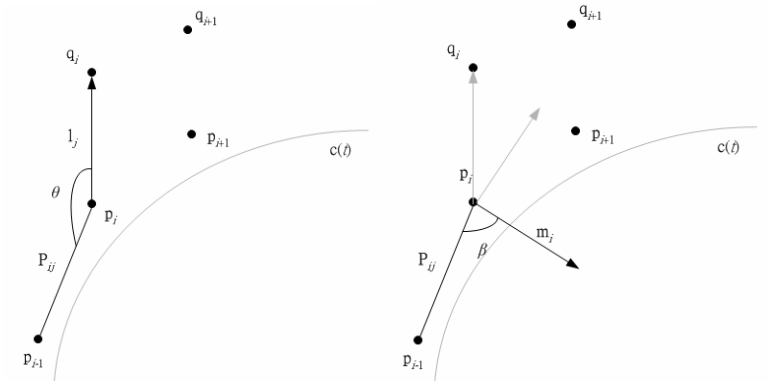
$$\mathbf{F}^{\theta}_{j} = K^{\theta} (\Delta\theta) \, \mathbf{n}_{\theta,i} * \mathbf{P}_{ij} \tag{8}$$

The same holds for $\mathbf{F}^{\lambda}_{i}$ and $\mathbf{F}^{\beta}_{i}$. A restoring force is applied to $\mathbf{p}_j$ in order to balance the torque. That is $\mathbf{F}^{\theta}_{i} = -\mathbf{F}^{\theta}_{j}$. For the twisting deformation on $\mathbf{P}_{ij}$, the torque at the point mass $\mathbf{p}_i$ is proportional to the angular displacement $\Delta\omega$. The torque can be expressed as:

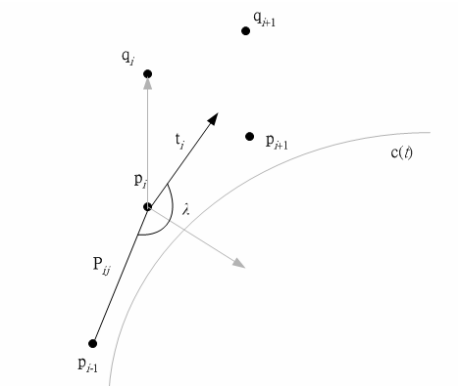
$$\tau^{\omega}_{i} = K^{\omega} (\Delta\omega) \, \mathbf{t}_i \tag{9}$$

where K^{ω} is the angular stiffness, $\mathbf{t}_i$ is the tangent vector at $\mathbf{p}_i$, and $\Delta\omega = \omega_i - \omega_j$. The torque is then converted to virtual forces on the point mass $\mathbf{q}_i$ and $\mathbf{q}_j$ of the orientation curve and they are expressed as:

$$\begin{aligned} \mathbf{F}^{\omega}_{i} &= K^{\omega} (\Delta\omega) \, \mathbf{t}_i * \mathbf{l}_i \\ \mathbf{F}^{\omega}_{j} &= -K^{\omega} (\Delta\omega) \, \mathbf{t}_j * \mathbf{l}_j \end{aligned} \tag{10}$$



A. Rotation angles (θ and β) of a curve-pair



B. Rotation angles (λ) of a curve-pair

Fig. 3.

3 Implementation

The proposed technique is implemented on a P4 3.0GHz PC with a GeForce 6600 256MB display card using VC++ and OpenGL. Figure 4 shows an example of simulating the cloth movement of a game character. Dynamic axial curve-pair is used to generate a prescribed cloth animation which can be further edited to produce a desired effect. Figure 5 shows an example of simulating the movement of a ribbon in a ribbon dance. A frame rate of 52 frames per second is achieved.

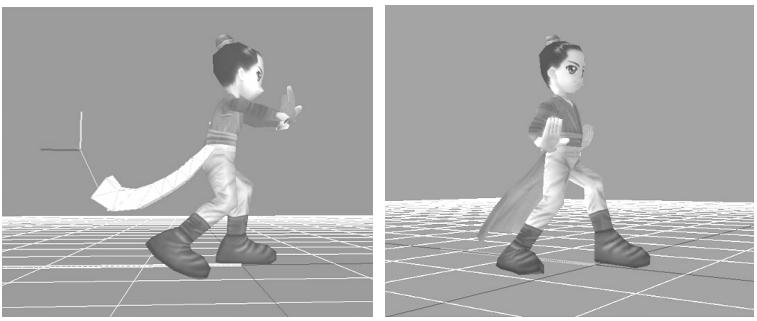


Fig. 4. Cloth simulation

Fig. 5. Simulating ribbon movement in a ribbon dance

4 Conclusion

We have proposed and implemented a framework for physics-based axial deformation. Our approach extends the axial curve-pair deformation technique to incorporate elastic deformation. The technique provides a solution to a problem in axial deformation where the lack of control on the local coordinate frame may lead to unexpected twist of

the object. A mass-spring model is adopted to implement a physics-based axial curve-pair skeleton. Interpolating a curve-pair to the point masses of the mass spring system emulates the physical deformation properties, and at the same time, ensures a smooth deformation of the object. By using special torsional spring elements to connect the local coordinate frames at a set of mass points approximating an axial curve, a mass-spring system is built on the 1-dimensional curve-pair structure with smooth bending and twisting behaviors. Physics-based deformation is performed on the point masses thereby deforming the embedded curve-pair. The proposed technique is useful for animating soft object such as ribbons and long cloth.

Acknowledgments. The work described in this paper was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region. (Project no. CUHK412106) and a Direct Grant (Project no. 2050413) from the Chinese University of Hong Kong.

References

1. Sederberg, T.W., Parry, S.R.: Free-form deformation of solid geometric models. In: ACM Computer Graphics (SIGGRAPH 1986) (1986)
2. Chadwick, J.E., Haumann, D.R., Parent, R.E.: Layered construction of deformable animated characters. Proceedings of ACM SIGGRAPH, Computer Graphics 23(3), 243–252 (1989)
3. Faloutsos, P., van de Panne, M., Terzopoulos, D.: Dynamic Animation Synthesis with Free-Form Deformations. IEEE Transactions on Visualization and Computer Graphics (1997)
4. Feng, J.Q., Ma, L.Z., Peng, Q.S.: A New Free-form Deformation through the Control of Parametric Surfaces. Computers&Graphics 20(4), 531–539 (1996)
5. Borrel, P.: Simple constrained deformations for geometric modeling and interactive design. ACM Transactions on Computer Graphics 13(2), 137–155 (1994)
6. Terzoppoulos, D., Qin, H.: Dynamic NURBS with geometric constraints for interactive sculpting. ACM Transactions on Graphics 13(2), 103–136 (1994)
7. Lazarus, F., Conquillart, S., Jancene, P.: Axial deformations: an intuitive deformation technique. Computer-Aided Design 26(8), 603–617 (1994)
8. Hui, K.C.: Free-form design using axial curve-pairs. Computer-Aided Design 34, 583–595 (2002)
9. Christensen, J., Marks, J., Ngo, J.T.: Automatic motion synthesis for 3D mass-spring models, Tech.Rep., MERL TR95-01 (1995)
10. Qin, H., Terzopoulos, D.: D-NURBS: A physics-based framework for geometric design. IEEE Transactions on Visualization and Computer Graphics 2(1) (March 1996)
11. Volino, P., Magnenat-Thalmann, N.: Comparing Efficiency of Integration Methods for Cloth Animation. In: Proceedings of Computer Graphics International 2001, Hong-Kong, China, pp. 265–274 (2001)
12. Choi, K.-J., Ko, H.-S.: Stable but responsive cloth. In: Proceedings SIGGRAPH 2002, San Antonio, TX, USA, pp. 604–611 (2002)
13. Baraff, D., Witkin, A., Kass, M.: Untangling Cloth. In: Proceedings SIGGRAPH 2003, San Diego, CA, USA, pp. 862–870 (2003)
14. Eberhardt, B., Weber, A., Strasser, W.: A Fast Flexible Particle-System Model for Cloth Draping. IEEE Computer Graphics and Applications 16(5), 52–59 (1996)

15. Haumann, D.R., Wejchert, J., Arya, K., Bacon, B.: An application of motion design and control in physically-based animation. In: Proceeding of Graphics Interface 1991, pp. 279–286 (1991)
16. Peng, Q.H., Jin, X.G., Feng, J.Q.: Arc-Length-Based Axial Deformation and Length Preserving Deformation. In: Computer Animation 1997, pp. 86–92. IEEE Computer Society, Geneva (1997)
17. Faux, I.D., Pratt, M.J.: Computational geometry for design and manufacturing. Wiley, Chichester (1979)
18. Klok, F.: Two moving coordinate frames for sweeping along a 3D trajectory. Computer-Aided Geometric Design 3, 217–229 (1986)
19. Lossing, D.L., Eshleman, A.L.: Planning a common data base for engineering and manufacturing. In: SHARE XLIII, Chicago (August 1974)
20. Fehlberg, E.: Low-order classical Runge-Kutta formulas with step size control and their application to some heat transfer problems, NASA Technical Report 315 (1969)
21. Cash, J.R., Karp, A.H.: A variable order Runge–Kutta method for initial value problems with rapidly varying right-hand sides. ACM Trans Math Software 16, 201–222 (1990)

3D Freehand Canvas*

Miao Wang, Guangzheng Fei, Zijun Xin, Yi Zheng, and Xin Li

Computers and Software School, Communication University of China, China
Animation School, Communication University of China, China
wm_cindy@163.com, gzfei@cuc.edu.cn, linus1115@gmail.com

Abstract. This paper presents a 3D freehand sketching system. Replacing the traditional 3D cartoon process, which contains modeling, texture mapping, it uses 2D input for generating projective strokes on a user-definable 3D canvas, which makes it possible that artist can sketch freehand in 3D space with no limitation. The freehand style animation could be created, which is nearly impossible for modern 3D animation tools. Most of the 2D painting features are also supported such as varieties of colors and abundant user-defined alpha stroke, and all of these features could be used in 3D space as well. Moreover, the 3D animation features could also be used, including space view, perspective view and freely 3D movement both for the object and the camera. It is a brand new tool both for traditional 2D cartoon makers and modern 3D animation craftsmen.

Keywords: 3D freehand, stroke, object animation, camera animation.

1 Introduction

Modern 3D animation movies are usually created through the following complicated steps, including modeling, texture mapping and varieties of methods of rendering with professional 3D animation software systems, which are exemplified by Autodesk 3D MAX and Maya. The whole process must be done by a group of well-trained craftsmen who are familiar with the complicated software and have a lot of experience on computer animation work.

Some features of the traditional 2D animations, however, sink into oblivion because of the emergence of the 3D animation, taking example as the freehand sketching style. Additionally, many excellent 2D cartoon designers still stand far apart from the 3D computer animation for the strange operations on modeling with a complex 3D software system. They prefer freehand sketching style even giving up the convenient 3D features. So they choose the tools like Painter or Adobe Flash to create styled 2D cartoons. The flaws can't be avoided either, which are that Painter supports varieties of stroke with no animation functions, and Flash provides abundant animation functions while it poorly supports the bitmap image, that means the lack of strokes and styles.

* This work was supported by a grant from NSFC (No. 60403037).

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 602–612, 2008.

Moreover the workload on dealing with those 2D animation tools is obviously heavier than that on handling 3D animation tools.

3D Freehand Canvas combines the advantages of 2D and 3D system. It is based on lines with sketching style which is considered as basic element of constituting an object, and includes translation of viewpoint for generating shielding effect. It provides a novel fashion for artists to create artistic work of 3D freely. And it is also software that is used to make 3D storyboard for supporting freehand sketching in the 3D space and animation creation. It is a convenient tool for artists who prefer creating artistic 3D works by using freehand drawing to doing them by modeling. Artists could get a "What you see is what you get" result. Additionally, 3D Freehand Canvas supports some animation functions which allows the objects and camera do some given movement.

2 Related Work

Freehand sketching plays an important role in both the past and the present phases of the design process. Because of being extremely intuitive mode to by using pencils and paper and they are always accessible, Most of the artists are inclined to drawing by hand with them. Sketchy presentation [10] could help the designers to communicate with each other better; the cognitive processes of the designer [7] can be expressed clearly though sketching. The goal of the research is to transform the way of freehand drawing to 3D data.

Some related work has been done. One approach to this is the usage of some 3D input devices, such as tracking devices, data gloves or armatures which are used in computer animation widely, there are many systems which has used 3D input devices for sketching, for example, 3-Draw [8], 3DM [2], HoloSketch [4] and the system presented by Diehl et al. [5]. Other systems which is for instance Tractus[9] and 3D6B editor[6] use the normal input devices such as mouse and canvas. Tractus supports a drawing canvas to do vertical movement, which makes the device able to maintain direct spatial mapping between real and virtual spaces. The method how Tractus allows drawing of non-planar curvature is remarkable. When drawing with a pen on the canvas surface and moving the canvas vertically, complex 3D paths can be constructed by user, which are difficult to input with 2D methods. 3D6B editor uses 2D input for generating projective strokes on a user-definable 3D grid. The rendering is done with a line-based renderer, and the data is never converted to 3D surface models but rather stored as strokes. The systems above achieve some functions of 3D freehand sketching. However, the shortage is obvious. Both of the systems deal with the simple lines without any styles and strokes. This is disadvantageous for artists to do the artistic works. Additionally, the drawing function is too limited to go into the realistic animation process. At last, they don’t have any animation functions.

Another path in the research is to use the strokes or paths directly without trying to convert them to 3D models. Cohen et al. [3] introduces a system where a 3D curve is defined by first drawing its screen plane projection and then its shadow on the floor plane. And Tolba et al. [11] present a system in which sketches of 3D space with fixed

camera position can be made. However, these results are only panoramic sketches in the whole scene, not full 3D with six degrees of freedom.

3D Freehand Canvas is the software that is used for freehand sketching and animation producing. It is a convenient tool for artists who prefer to freehand drawing to create the artistic 3D works. Additionally, 3D Freehand Canvas supports some animation functions which allows the objects and camera do some given movement. The traditional storyboard features are extended.

3 System Structure

The 3D Freehand Canvas System is a sketching system for drawing non-precise 3D sketching and producing sketching style animation. It uses 2D input for generating projective strokes on a user-definable 3D grid. The rendering is done with the OpenGL render, and the data is never converted to 3D surface models but rather stored as strokes.

Because of the differences with the traditional rendering method of 2D and 3D, our system requires to position lines to 3D space for rendering, so our system is bound to set canvas in 3D space, which is a kind of controllable grid canvas on which strokes can be drawn freely, and it can be manipulated with full six degrees of freedom, and lines will be projected onto the canvas, thus the information of lines can be stored in canvas. Meanwhile, a serial of canvas which have lines information constitute an object, every object has its own axis, when drawing an object, canvases and lines belonging to the object will be translated into the object coordinates system, editing axis of object means to edit object.

The movement and rotation of canvas play an important role for drawing an object freely in 3D space. The relationship between canvas and object in 3D space may be dramatically complex, which includes that canvas is operated as the object holding its position in the world coordinates system or the opposite; or after editing object, canvas which is included in current object should also be edited, in such cases, canvas and object must be tied by some special relationship, so just one canvas is not enough, our method is to establish two kind of canvas. One is temporary canvas, and the other one is real canvas. The former is used to show current canvas, it does not store any stroke. The latter one which will be constructed after the first stroke on the canvas is finished will store all the strokes drawing on the canvas, and it will be appended to the current object. Every temporary canvas has a flag, when the flag is equal to 0, it means real canvas has existed and user will draw on this real canvas, it is no need to construct a new real canvas; when the flag is equal to 1, it means current object has moved or rotated, if user continues drawing, the new stroke requires a new real canvas, the algorithm will construct a new real canvas which is the same as the current temporary canvas. When current canvas or current object moves or rotates, the flag must be equal to 1.

Our system also provides non-planar canvas. The grid of this canvas is not limited to a plane any longer; it can be expressed in different kinds of sharps according to the sharp of stroke drawing by users. Non-planar canvas is generated by sampling current

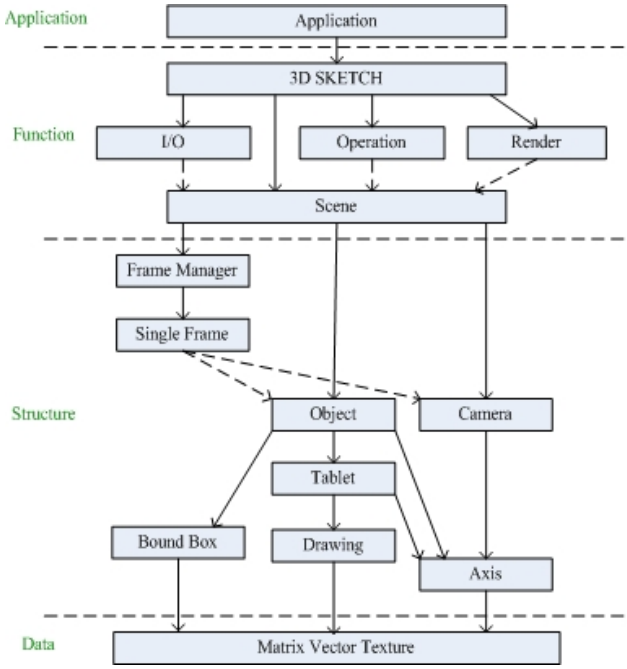


Fig. 1. System structure

stroke, firstly, calculating the arc length of the stroke by Gaussian quadrature, then establish adaptable subdivision table of the stroke, the sampling points are the points on the stroke which have same interval. The distance between two sampling points is the width of non-planar mesh.

In short, the structure of our system is that there is a scene including serials of objects and a camera, and an object includes several canvases which also include many strokes, the camera and every object can be set key frames animation (figure 1).

4 Stroke Model

One of the significant differences between other system and ours is that our system defines a kind of stroke with sketching style, which contributes to constructing artistic work. In this section, I describe how to generate the stroke model that I have used.

4.1 Stroke Mesh

In our system, the generation of our stroke is similar to the method of Stroke-Based Rendering [1]. A basic brush stroke is defined by stroke curve, stroke mesh, stroke color, stroke texture. The curve is also an endpoint-interpolating cubic B-Spline defined by a set of control points. Curve points can be computed by recursive

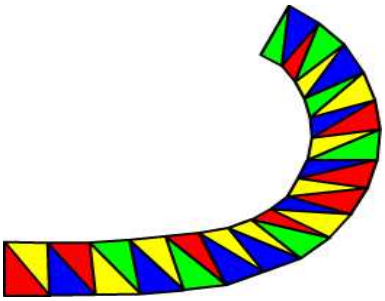


Fig. 2. Stroke mesh

subdivision. But the difference is that our stroke must be drawn on canvas; it means that each stroke needs to be converted to 3D space, so the first step is to calculate the points that intersect on current canvas, and these points are the control points of a stroke. And then, I bound to convert a stroke into a stroke mesh after all the control points have been inputted. The basic technique for generating stroke mesh is to tessellate the stroke into a triangle strip.

Considering that a dense list of control points P_i and a brush thickness R should be moderate, we can get the geometrical topology of the stroke by the following steps (Figure 2): curve tangents should be computed at each control point firstly and then is the normal directions of curve. An adequate approximation to the tangent for an interior stroke point P_i is given by $V_i = V_{i+1} - V_{i-1}$. The first and the last tangents are $V_0 = P_1 - P_0$ and $V_{n-1} = P_{n-1} - P_{n-2}$. The next step is to calculate the points on the boundary of the stroke, according to the distance R, these points will offset by R along the normal direction. So the geometrical topology of the stroke can be generated as the Figure 2 showing above.

When the topological structure has been finished, it means that a stroke has become a triangular mesh, so there are two ways to show a stroke, one is to fill the mesh with a single color, and the other way is to use texture mapping.

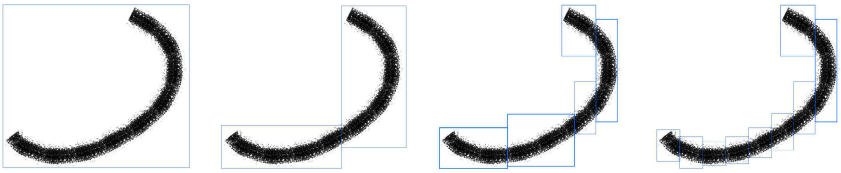
For expressing abundant stroke with sketching style, our system supports to texture mapping, TGA texture with alpha value can be considered as stroke texture.

4.2 Storage Optimization of Stroke

The number of points on stroke mesh is dramatically large generated by the method above, it can make stroke look smooth. But when the objects are stored in a file, we need to record the stroke mesh and texture, the size of file will be large by this way; or if the object is far away from the camera, there is no need to give a very precise expression of the shape of the stroke, approximate expression is enough, thus, the storage of stroke should be optimized, namely, the number of points should be decreased in order to economize space of storage.

Our method is when a stroke finished, finding the feature points on the stroke firstly, our arithmetic adopt a method which makes the spatial curve parameterized by arc

length, calculates the adaptive subdivision table by given a desired accuracy, the spatial coordinates of these points calculated by the adaptive subdivision table are the feature points, every two feature points has a bounding box, namely feature bounding box. Then, getting the bounding box of the whole stroke, the stroke in its bounding box will be converted to a new texture, then get the texture coordinates of every bounding box on the texture, the alpha value of the texture is equal to 0, if on which there is no stroke, otherwise, the alpha value is equal to 1. When a stroke is saved, only the feature bounding boxes of the strokes and a texture are stored. (Figure 3) In this way, not only the storage space is optimized, but also efficiency is improved.



(a) (b) (c) (d)

Fig. 3. When given different desired accuracy, the key bounding boxes will change, the bigger the desired accuracy is, the more the number of the feature bounding boxes are

4.3 Stroke Layer

In OpenGL environment, because of the error which is caused by float computing, the cross strokes on the same canvas presents a "sieve" in the cross area (Figure 4(a)). The reason is that users ideally consider that strokes on the same canvas has the same z-depth value while they are differ at random, which is changing in real time with the different view.

Our solution to this problem is to mark stroke with an integer orderly when it is drawn on the canvas, then add a tiny value to the z-depth value of the stroke with the bigger mark value when it is crossed with another one. (Figure 4(b))

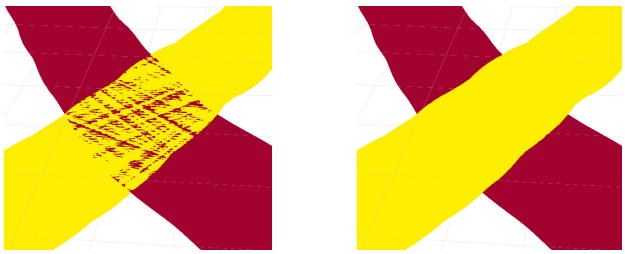


Fig. 4. (left) a "sieve" in the cross area because of the error which is caused by float computing. (right) "Sieve" disappears by adding a tiny value to the z-depth value of the stroke.

To judge whether the two strokes are crossed or not, it is important to find the feature point in the curves. After computing the feature points by using the adaptive subdivision table, the curve could be described approximately with the polygonal lines that consist of the feature points orderly. Then test the cross among the polygonal lines instead of the mesh stroke.

4.4 Tiling Texture

The amounts of the textures depend on the curve length. The length of every stroke curve is computed after the operation of adding control points works, which is the sum of the distance of every two sequential control points. The quotient of the curve length divided by the width of a texture is defined as the amount of the textures, which is used in texture mapping.

According to the texture mapping method of OpenGL, texture coordinates are described as the value that starts at 0 and ends at 1 for one texture. The number above 1 will be considered to use another same texture. So the texture mapping of a stroke can be generated by given the amount of the textures, the texture will be mapped to stroke mesh one by one, the quality and size of texture can be ensured by this way, and the texture will not be stretched.

At present, a stroke composes of stroke mesh and texture mapping, in the future, we will adopt procedural texture, it will generate the beginning and the end of a stroke automatically, in addition, procedural texture will be affected by the pressure and speed in the process of drawing, and the direction of the brush, the stroke with sketching style generated by this way, will be better than it by stroke mesh.

5 Animation

It is difficult for traditional 2D animation to express the movement of camera and objects clearly, while our system provides two kinds of animation forms, one is object animation, and the other is camera animation. Both of the two animation forms adopt key-frame animation. Each key frame records key frame index, the axis information of current object and speed curve. Our system also supports users to produce rigid movement of multi-objects and the camera simultaneously to make users find satisfaction in setting objects and camera animation.

Key-frame animation is familiar with other animation software, but in traditional 2D animation software, most of them, for example Flash, adopt a method for generating animation that the motion path should be drawn first, and make objects attach themselves to the points on the path, it means that needs to make objects be bound to path given, and then, by setting key frames to calculate the middle position of each frame where the object should be on the path. This method, however, doesn't fit for paths in 3D space, in that it is very difficult and discommodious for tying the position of the object to the points of path, so in our system, it uses a method that it is no need to

draw a path and bind objects to it, the path generated is by getting the key position of the object in information of each key frames, according to the list of the key position of object, when all key frames has finished setting, the path will be generated automatically as a form of Catmull-Rom Spline curve. By using this method, it is convenient to draw a path in 3D space, without any binding operations.

After generating a path curve, the speed curve can be generated automatically too. Each object has its own motion, and each motion has one path and one speed curve. In our system, speed curve adopts line curve mode and Multi-Hermite curve mode. Speed curve, actually, is a distance-time curve. Both distance and time can be obtained by the position of the object in 3D space and specifying key frames. So the most important thing of generating a speed curve is to calculate tangents of these control points. Our method is to calculate tangents based on control points given. The tangents of two endpoints are calculated firstly, the directions of them are along the direction of joining of the two endpoints and in opposite direction. And then, the tangents of middle points can be obtained by the previous point and posterior point of this one, so find these two points which the mid-one should be insert, and the direction of joining of them is the tangent of the mid-points.

6 Experimental Results

Our system can create model on a standard PC at interactive rates, user can draw objects and set object and camera animation. User can also change the sketched style of stroke by selecting different stroke files.

In our system, we adopt a method of drawing the outline of object to represent an object in 3D space, and the process of drawing is a kind of mode of “what you see is what you get”. The system can simulate the sketching style of artists, describe object by using strokes with different kinds of sketching style, and the whole scene can be rendered in real-time. The system can generate not only the doodle effect like drawing on 2D canvas, but also the shielding effect between objects. Thus, our system has significant advantages compared with 2D and 3D software: compared with 2D system, viewpoint in our system will be changed and the shielding relationship between objects will be also changed following the viewpoint, figure 5 shows the shielding and perspective effect of different objects from different viewpoints; compared with 3D system, our system can describe objects with sketching style, which is nearly impossible in 3D software, an object composes of just some strokes without modeling, the system is much freer than 3D software, and it is in line with the habits of artists, has dramatic performance of art, figure 5(b) is different from the others, the sketching style of strokes in these three pictures looks like crayon, and the others simulate the style of watercolor painting, these pictures represent different drawing styles. The following ones (Figure 5) show single objects with changing viewpoint, (Figure 5(a)) and the whole scene with fog around it which composes of several objects.(Figure 5(b),(c)).

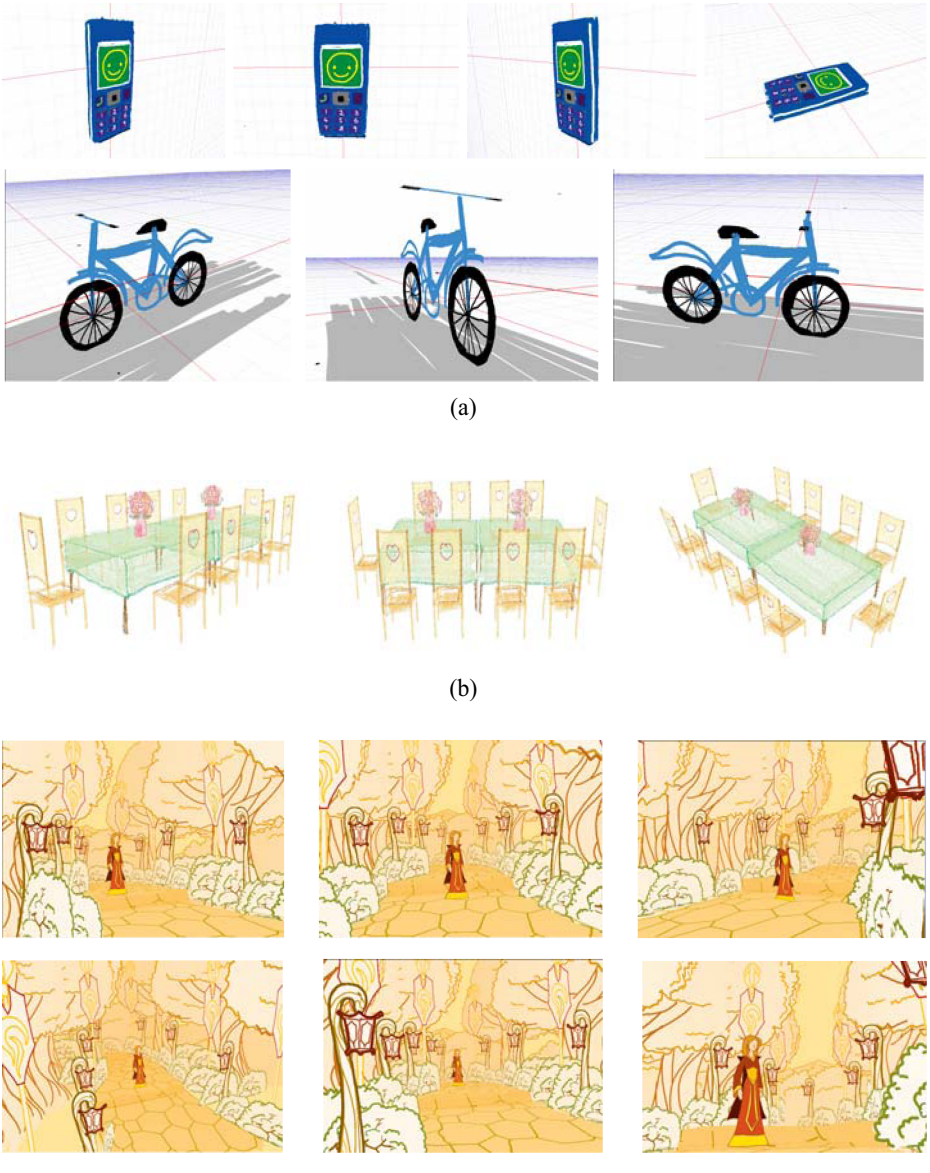


Fig. 5. Result pictures. (a) Showing the single objects from different view point. (b) These three pictures are a little more complicated than the ones in (a), which composes of multi-object, and the drawing style is crayon while others are watercolour. (c) Showing a whole scene which describes nightfall of magic world with light yellow fog around it, it is more complex than others, and includes the movement of people and viewpoint.

7 Conclusion and Future Work

In this paper we present an approach for establishing strokes with sketching style, drawing objects and generating animation. The system is able to create rather complicated stroke mesh with almost only sketch input, which can be used to generate sketching style by specifying stroke texture and appending texture to mesh. Apart from defining stroke mesh, our system also allows using stroke texture with Alpha value, and does storage optimization of stroke. The system runs at interactive rates on a standard PC. The whole system is adapted well to non-expert users.

There would be also lots of works to do to enrich the functions of this software. Firstly, we provide some sketching styles of stroke to users, but the variety of stroke style is still not rich, in the further study, we intend to map multi-texture to stroke mesh, and allow users define stroke style by themselves. Secondly, strokes can be edited by user, such as movement, rotation, copy, paste, but can not transform, in the further work, the sharp of stroke will be transformed by pushing and pulling. Third, because of the limit of the drawing styles, just some transparent texture created by other software or some common format could be used in the storyboard; extending brushwork library will be an effective supplement for the 3D Freehand Sketching system. Fourth, the system is able to express a scene containing multi-object rendering with depth test, but objects at distant position is too clear to have an effect on near ones, thus, an effect of distant objects disappearing gradually is necessary for users, which makes near ones more clearly.

References

1. Hertzmann, A.: Stroke-Based Rendering. In: SIGGRAPH 2002, pp. 3.1-3.31 (2002)
2. Butterworth, J., Davidson, A., Hench, S., Olano, M.T.: 3DM: a three dimensional modeler using a head-mounted display. In: SI3D 1992: Proceedings of the 1992 symposium on Interactive 3D graphics, pp. 135–138. ACM Press, New York (1992)
3. Cohen, J.M., Markosian, L., Zeleznik, R.C., Hughes, J.F., Barzel, R.: An interface for sketching 3D curves. In: SI3D 1999: Proceedings of the 1999 symposium on Interactive 3D graphics, pp. 17–21. ACM Press, New York (1999)
4. Deering, M.F.: HoloSketch: a virtual reality sketching/animation tool. ACM Trans. Comput.-Hum. Interact. 2(3), 220–238 (1995)
5. Diehl, H., Müller, F., Lindemann, U.: From raw 3D-sketches to exact CAD product models —concept for an assistant-system. In: Hughes, J.F., Jorge, J.A. (eds.) Eurographics Workshop on Sketch-Based Interfaces and Modeling. Eurographics (August 2004)
6. Kallio, K.: 3D6B Editor: Projective 3D Sketching with Line-Based Rendering. In: Igarashi, T., Jorge, J.A. (eds.) Eurographics Workshop on Sketch-Based Interfaces and Modeling (2005)
7. Lim, C.-K.: An insight into the freedom of using a pen: Pen-based system and pen-and-paper. In: Proc. 6th Asian Design International Conference (October 2003)
8. Sachs, E., Roberts, A., Stoops, D.: 3-Draw: A tool for designing 3D shapes. IEEE Comput. Graph. Appl. 11(6), 18–26 (1991)

9. Sharlin, E., Sousa, M.C.: Drawing in space using the 3D tractus. In: 2nd IEEE Workshop on New Directions in 3D User Interfaces (IEEE VR 2005) (March 2005)
10. Schumann, J., Strothotte, T., Laser, S., Raab, A.: Assessing the effect of non-photorealistic rendered images in CAD. In: CHI 1996: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 35–41. ACM Press, New York (1996)
11. Tolba, O., Dorsey, J., Mcmillan, L.: Sketching with projective 2D strokes. In: UIST 1999: Proceedings of the 12th annual ACM symposium on User interface software and technology, pp. 149–157. ACM Press, New York (1999)

Sparse Key Points Controlled Animation for Individual Face Model

Jian Yao, Yangsheng Wang, and Bin Ding

Institute of Automation, Chinese Academy of Sciences, Beijing, China, 100080
jian.yao@ia.ac.cn

Abstract. We use RBF deformation and normal projection to regulate the scanned facial model, which makes their topology equivalent to a regular grid mesh and can generate principle components. Then the synthesized individual face model can be directly flatten to a regular plan, so the motion vectors of vertexes can be interpolated with barycentric coordinate. The regulation and animation remapping needs less than 40 key points and can work in real-time

Keywords: Sparse key points, Animation remapping, Regular mesh, Barycentric coordinate.

1 Introduction

Individual facial modeling and animation has a wide applicability in entertainment animation, interactive games, HCI, telepresence and medical research. And it's always a challenging and interesting task. Also there have been many excellent methods proposed for two main steps of the problem, modeling and animation.

3D morphable model[1] performs very well for individual facial modeling from one or more photos. The method applies PCA (principal component analysis) on 3D coordinates and colors of face model, then any new face could be represented by linear combination of these components. In the pretreatment of regulation, it computes correspondence between two faces with an optic flow algorithm, which performs well but cannot modify the density of various region. For instance, the eye region may need more detailed information than the cheek region, which means its better to be denser for the vertexes in eye region. Also the symmetry of vertex cannot be guaranteed.

For generating animation of face model, expression cloning[2] performs well and can remap predefined morph animation to a target model with any topology. This method computes dense correspondence for each vertex in source and target models and uses this information to achieve motion remapping. Expression cloning requires the source control points to be as dense as possible, which makes it difficult for acquisition of source motion data.

To overcome the limitations above, we introduce regular grid mesh and normal projection to regulate scanned face model, which makes PCA mesh more unique and the dense of vertex are uniquely determined by location of key points.

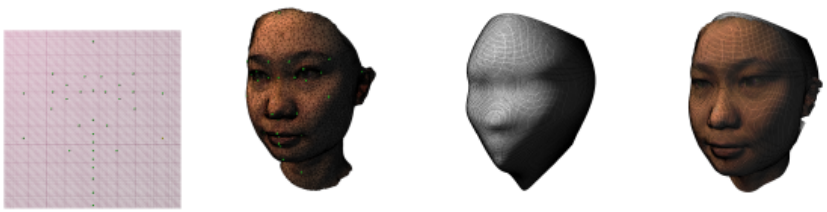
Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 613–618, 2008.

Also the regulated model can be directly flattened. After remapping of key points' motion vector, we interpolate vertexes' motion vector with barycentric coordinate. All the process needs less than 40 key points.

The rest of the paper is organized as follows. Section 2 describes how to regulate scanned face models, especially about normal projection and the accelerating with octree. Section 3 describes how to remapping animation from motion capture data to the regular face model. Section 4 includes postprocess about mouth division. Then Section 5 shows some examples and draws a conclusion.

2 Regulation of Scanned Models

In PCA based method of individual face generation, the regulation of scanned face model is an inevitable step. The regulation means to remesh all the face models to the same topology and make sure any one vertex has the same anatomy meaning for every model. Also the uniformity of vertexes should be taken into account, since sparse mesh will reduce the quality of texture. We use RBF and normal projection to regulate scanned face models(Fig.1). Also, the vertexes' color of regular model is retrieved during normal projection process.



(a) Regular grid mesh(128*128), green points represent key points (b) Scanned face model (c) After RBF deformation (d) After normal projection and color retrieval

Fig. 1. Regulation of face model

RBF (Radial Basis Function) is a supervised neural network and a powerful tool for geometry smoothing and deformation [3]. RBF could perform deformation between any two meshes with correspondence key points. Its performance is determined by position and amount of the key points. Fig.1(c) shows the result of RBF deformation. The RBF equation is

$$f(p_i) = \sum_{j=1}^{n} \omega_j \cdot h_j(p_i) \tag{1}$$

where $p = (x_i, y_i, z_i)^T$ and $f(p_i) = (x_i^{'}, y_i^{'}, z_i^{'})^T$ is input and output coordinates respectively. ωj is weight coefficient to be computed. And here is the basis function

$$h_j(p_i) = \sqrt{\left\| P_i - P_j \right\|^2 + s_j^{\,2}} \tag{2}$$

where $s_i = \min_{j \neq i} \left\| p_i - p_j \right\|$ A regularization parameter λ is used to minimize the cost function

$$C(\omega) = e^T e + \lambda \cdot \omega^T \omega \tag{3}$$

where e is the error vector between the training input and output coordinates.

After RBF deformation, the shape of source grid mesh is roughly similar to the target model but the detailed difference is very large. A projection process must be applied to make the source mesh "close-fitting" to the target model(Fig.1(d)). J. Noh[2] use cylindrical projection to improve the RBF deformation, but it could have large distortion if the target triangle face is close to horizontal. We use the normal vector of vertex as the direction of projection. This makes the projection result more uniform. The comparation can be seen in Fig.2.

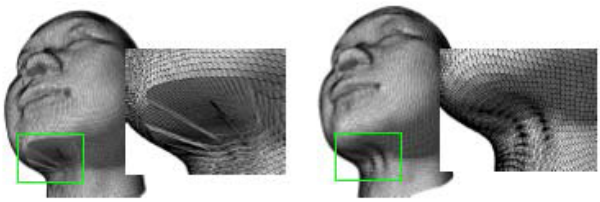


Fig. 2. Comparision of cylindrical projection (left) and normal projection (right)

The projection process needs to traverse all the vertexes of source mesh and all the triangle faces of target model, so this process is time-consuming. We introduce octree to accelerate it(Fig.3). For each vertex, the coordinate and normal vector define a line in 3D space (green line). We compute to determine the octree nodes that intersect with the line, then only the triangle faces (red region) in intersected nodes will be taken into account. This will commonly spare more than 75 percent computation time. All the intersection points will be computed and the nearest to the vertex is chosen as projection point.

The scanned models are acquired with 3D scanner and contain color information. Consequently, the regular face model also needs color to render. Vertex of grid mesh is now located on one triangle face of the scanned model. Thus we can get the barycentric coordinate and use it to compute vertex color(Eq.4).

$$\tilde{C}_g = \alpha \cdot \tilde{C}_{s0} + \beta \cdot \tilde{C}_{s1} + \gamma \cdot \tilde{C}_{s2} \tag{4}$$

where $\tilde{C}_g = (R_g, G_g, B_g)$ is the color of grid mesh vertex, (α, β, λ) is barycentric coordinate, and $\tilde{C}_{g0}, \tilde{C}_{g1}, \tilde{C}_{g2}$ is color of the three triangle vertexes respectively.

With all the process above, scanned face models have been regulated to a same regular topology. Then, we can construct a matrix with vertexes' coordinate as the

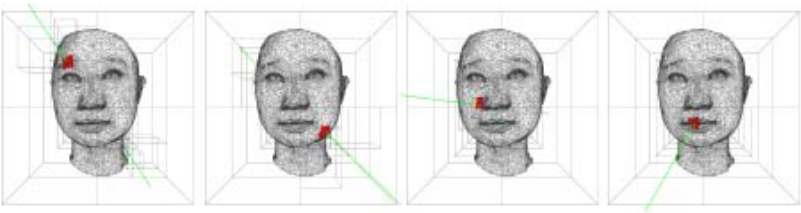


Fig. 3. Projection acceleration with octree

Fig. 4. Regular mean face model

column and employ eigen-decomposition to get principle components of the shape. Color components can be obtained with the same method. Various individual face shape and skin color can be represented by linear combination of these components.

3 Animation Interpolation

The source animation data is obtained from Motion Capture with only 39 markers. For the shapes of performer's face and grid model are different, so direct transfer of the motion data between models is inappropriate and may induce errors such as overlapping. The direction and magnitude of motion vectors need to be modified before remapping. We employ local bounding box[2] to do this job. Bounding box is a directional cuboid contains a vertex and all its neighbor vertexes. As shown in Fig.5, V and V' are key point of source and target models respectively, m is motion vector of V . The change of direction and magnitude from source to target bounding box determines the transform matrix R, thus $m' = R \cdot m$ is the motion vector of V'.

Our goal is realistic facial animation with sparse key points, which require an unique interpolation method. Parameterization is the optimal method to achieve unique correspondence and interpolation. Since all the face models have been regulated, which means the model can be directly flattened similarly as plane parameterization. When a vertex is flattened onto a plan, it will be located in one triangle face of a predefined key point mesh. Similar to the color retrieval, motion vector of the vertex is interpolated by the barycentric coordinate and motion vector of three key points. The interpolation equation is the same as Eq.4.

For bounding box is local property of the key mesh and varies with the topology, so it's inevitable for errors in motion vector remapping. We smooth motion vectors using

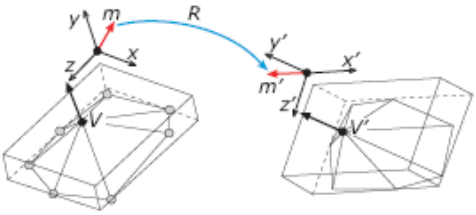


Fig. 5. Local bounding box

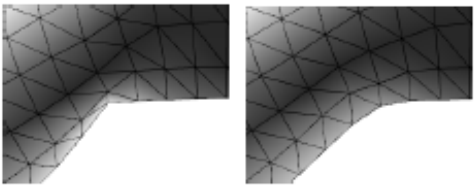


Fig. 6. Motion vector smooth

Eq.5, in $\vec{m}_j$ which is smoothed motion vector of vertex i, $\vec{m}_j$ is the motion vector of its neighbor vertex j, and $\omega_{ij} = \frac{1}{\|p_i - p_j\|}$ is reciprocal of the distance between vertex i and vertex j. Fig.6 shows an example, the inflexion in underlip is corrected.

$$\vec{m}_i' = \frac{1}{\sum_j \omega_{ij}} \cdot \sum_j \omega_{ij} \cdot \vec{m}_j' \tag{5}$$

4 Mouth Process

To enable the mouth to open, the mesh must be cut off along the mouth contact line and vertexes should have motion vectors in different directions. Otherwise, vertexes in different lips could have motion in same direction and thus the mouth cannot open. Since all the vertexes of lip contact line lie in one row of the grid mesh, we only need to duplicate these vertexes and assign them to triangle faces in different lips to achieve lip division.

Vertexes in lip region are affected by key points in both sides, which requires distinguishing upper and lower lip. We use Dijkstra's algorithm[4] to compute the shortest path from vertex in lip region and midpoint of upper lip. If the path intersects with lip contact line, the vertex should be marked as lower lip. Otherwise, no intersection means the upper lip vertex. Thus we can restrict the influence of lip key points only in vertexes of its own side. Fig.7 shows a comparison of the result with and without lip division. Yellow point is a key point and vertexes influenced by it are represented by red points, the transparency represents influence scale.

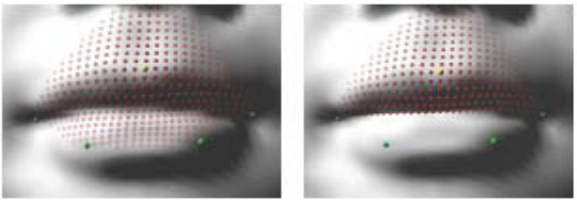


Fig. 7. Lip division

5 Result and Conclusion

Fig.8 shows some examples of individual face modeling and animation remapping. Each model has 16384 vertexes and 21292 triangle faces. The animation represents rational difference for different morphological character of the human face, and can performance more than 100 fps. This approach can satisfy common application such as games. And possible ways to improve it include using skeleton-muscle models and adding teeth or eye mesh to refine the detail.

References

[1] Blanz, V., Vetter, T.: A morphable model for the synthesis of 3D faces. In: Proceedings of the 26th Annual Conference on Computer Graphics and interactive Techniques, International Conference on Computer Graphics and Interactive Techniques, pp. 187–194. ACM Press/Addison-Wesley Publishing Co., New York (1999)

[2] Noh, J., Neumann, U.: Expression cloning. In Proceedings of the 28th Annual Conference on Computer Graphics and interactive Techniques. In: SIGGRAPH 2001, pp. 277–288. ACM, New York (2001)

[3] Pighin, F., Hecker, J., Lischinski, D., Szeliski, R., Salesin, D.H.: Synthesizing realistic facial expressions from photographs. In: Proceedings of the 25th Annual Conference on Computer Graphics and interactive Techniques, SIGGRAPH 1998, pp. 75–84. ACM, New York (1998)

[4] Johnson, D.B., Johnson, A.: A Note on Dijkstra's Shortest Path Algorithm. J. ACM 20(3), 385–388 (1973)

[5] Tutte, W.T.: Convex representations of graphs. Proc. London Math. Soc. (10), 304–320 (1960)

[6] Sheffer, A., Praun, E., Rose, K.: Mesh parameterization methods and their applications. Found. Trends. Comput. Graph. Vis. 2(2), 105–171 (2006)

[7] MathWorld, Wolfram Research, `http://mathworld.wolfram.com`

[8] Sumner, R.W., Popovic, J.: Deformation transfer for triangle meshes. In: SIGGRAPH 2004, pp. 399–405. ACM, New York (2004)

[9] Langer, T., Belyaev, A., Seidel, H.-P.: Spherical Barycentric Coordinates. In: Proceedings of the Fourth Eurographics Symposium on Geometry Processing (2006)

Networked Virtual Marionette Theater

Daisuke Ninomiya[1], Kohji Miyazaki[1], and Ryohei Nakatsu[1,2]

[1] Kwansei Gakuin University, School of Science and Technology
2-1 Gakuen, Sanda, 669-1337 Japan
{aaz61232,miyazaki,nakatsu}@kwansei.ac.jp
[2] National University of Singapore, Interactive & Digital Media Institute
Blk E3A #02-04, 7 Engineering Drive 1, Singapore 117574
idmdir@nus.edu.sg

Abstract. This paper describes a system that allows users to control virtual marionette characters based on computer graphics (CG marionette characters) with their hand and finger movements and thus perform a marionette theatrical play. The system consists of several subsystems, and each subsystem consists of a web camera and a PC. It can recognize a hand gesture of its user and transform it into a gesture of a CG marionette character. These subsystems are connected through the Internet, so they can exchange the information of the CG marionette character's movements at each subsystem and display the movements of all characters throughout the entire system. Accordingly, multiple users can join the networked virtual marionette theater and enjoy the marionette play together.

Keywords: Marionette, puppet, virtual theater, hand gesture, image recognition.

1 Introduction

The culture of controlling puppets with the hands to perform theatrical play has been common throughout the world from ancient times. In Japan, there is a type of puppet theater called Bunraku, which arose about three hundred years ago [1][2]. In Europe, too, various kinds of puppet play have been performed and enjoyed. The puppet play using a puppet called a "marionette" has been the most popular variety [3]. Marionette play and puppets have become very popular in recent years, largely due to the movie called "Strings [4]" (Fig. 1). This paper describes a networked virtual marionette theater that is basically a distributed system consisting of several subsystems connected through the Internet. Each subsystem can recognize the hand and finger gestures of the person in front of its web camera and then transform them into the motions of a marionette character based on computer graphics (CG marionettes). Each subsystem exchanges the information of actions performed by its marionette character with such information from the other subsystems. The display of each subsystem shows a virtual scene where multiple marionette characters, each controlled by a different user, interact. Thus multiple users, even if they are in separate locations, can gather in a virtual marionette theater and perform a theatrical marionette play.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 619–627, 2008.

Fig. 1. A scene of "Strings"

2 Related Works

Technologies based on three-dimensional computer graphics have made tremendous progress in recent years. We can see photographically real CG objects and CG characters in movies and games. Furthermore, the technologies based on CG animation have also progressed rapidly. Animations of fluid [5] and the destruction of objects [6] have been studied. Moreover, the movements of a crowd based on an artificial-intelligence approach [7] and movements of humans based on inverse kinematics [8] have been proposed. Motion capture systems have been widely used for the control of CG characters [9]. Although the realization of human-like motions of CG characters has been eagerly pursued, the realization of marionette-like motions has seldom been studied. Since the movements of marionette characters are unique and have been loved by people throughout history, it is worth studying a system by which non-experts of marionettes can easily manipulate their movements and generate marionette-like behaviors using CG characters.

3 System Concept

The following elements typically compose a marionette theater.

(1) Puppets called "marionettes"
(2) Speech of each puppet
(3) Scene settings
(4) Music

In a large performance, various kinds of marionette puppets appear and the scene settings are changed frequently, depending on the story's plot, and even a live orchestra is sometimes used to generate music. Therefore, even if people wanted to enjoy manipulating marionette puppets and creating theatrical play, it could be very difficult. On the other hand, if we introduced virtual marionette characters based on computer graphics instead of using real marionettes with physical bodies, it would

become significantly easier for users to generate and change most of the above elements of marionettes, speech, backgrounds, and music. In addition, by developing a networked virtual marionette theater, multiple users, manipulating their own marionette characters, can gather in a virtual theater and let their virtual puppets interact with other puppets, thus creating the performance of a virtual theatrical play.

4 System Structure

4.1 Overview

The entire system is made from a group of subsystems connected through a network. The structure of the whole system is shown in Fig. 2, and the structure of each subsystem is illustrated in Fig. 3. Each subsystem consists of a PC and a web camera

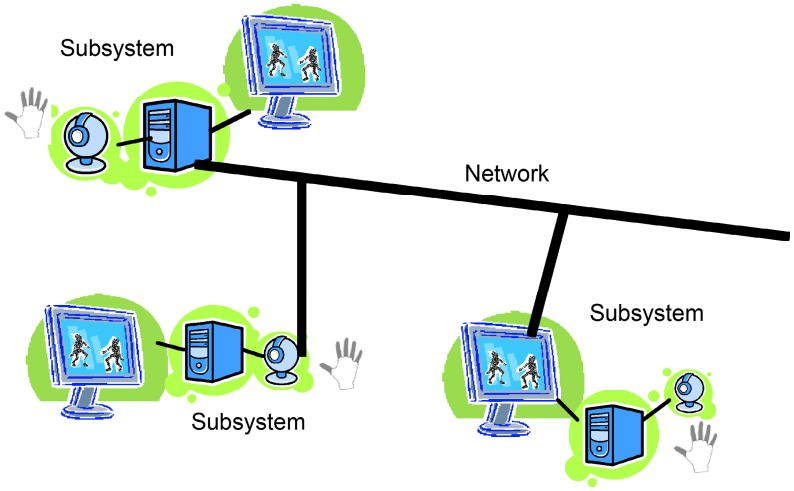


Fig. 2. Structure of entire system

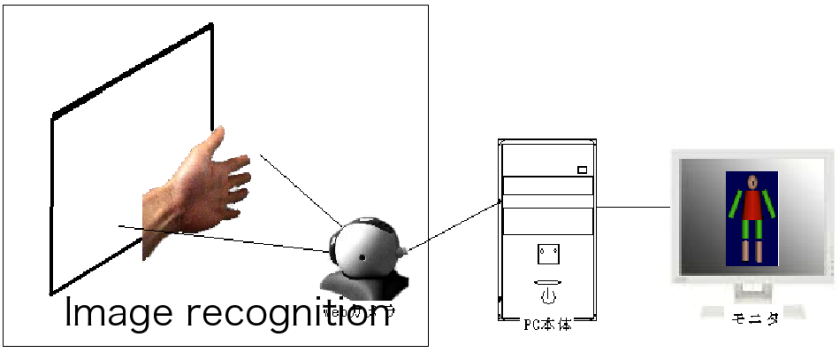


Fig. 3. Structure of subsystem

The posture of a user's hand is captured by the web camera, and then hand-gesture recognition is carried out. Then the recognition result of a hand posture is reflected in the gestures of a CG marionette character.

4.2 Hand-Gesture Recognition

In this section, a real-time hand-gesture recognition method is described for use in the recognition of a user's hand gesture for each subsystem [5]. There have been several research efforts on the real-time recognition of hand gestures [6][7]. Most of them use rather complicated systems such as multiple cameras. On the other hand, we tried to develop a simpler system using a single web camera. The recognition process consists of the following sub-processes.

4.2.1 Extraction of Hand Area (Fig. 4)

Using the color information of a hand, the image area corresponding to a hand is extracted from the background. In this case, HSV information obtained by the transformation of RGB information is used. Then, using a median filter, the noise contained in the extracted image is deleted.

Fig. 4. Extraction of hand area

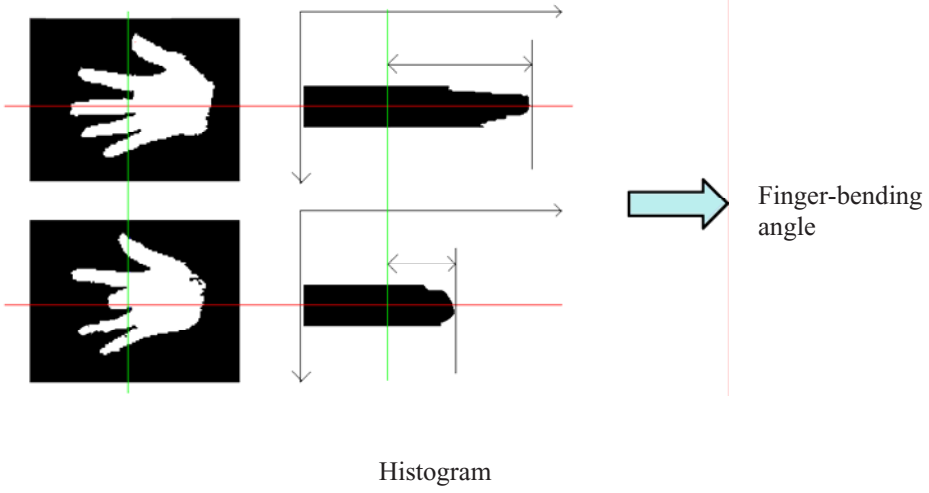


Fig. 5. Extraction of finger-length information

4.2.2 Extraction of Finger Information Using Histogram

The length of each finger is calculated by using simple histogram information. Figure 5 shows the information of a histogram corresponding to finger length. Depending on the angle of finger bending, the height of the histogram varies. This means that from the height information of the histogram, the bending angle of a finger can be calculated.

4.2.3 Optimization of Separating Each Finger's Histogram

Depending on the angle of each finger against the x axis (or y axis), it is sometimes difficult to clearly separate a histogram corresponding to each finger. Therefore, for the information extraction of each finger, rotation transformation is carried out to achieve the optimum separation of partial histograms corresponding to each finger.

4.2.4 Bending-Angle Estimation of Each Finger

Figure 5 also shows a comparison between two histograms varying with the bending angle of a finger. By comparing the length of a histogram to the original (longest) histogram when the bending angle is zero, the bending angle of the finger is calculated.

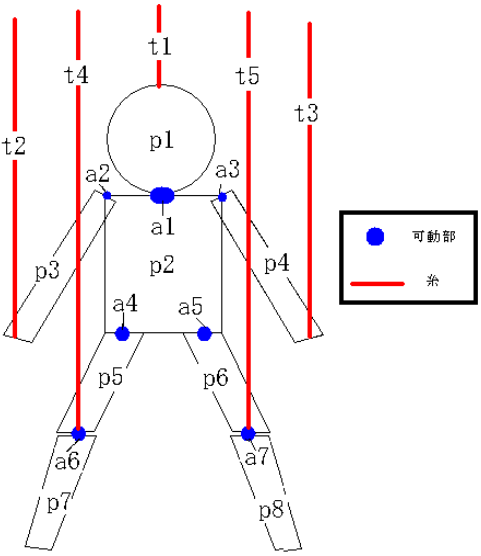


Fig. 6. Model of a virtual marionette

4.3 Control of CG Marionette

Each finger is assumed connected to a certain part of a CG marionette through a virtual string. The relationship between five strings and the part of the marionette to which each string is attached is illustrated in Fig. 6. Here, t1 ~ t5 are virtual stings, and p1 ~ p8 are the parts composing the marionette model, where a1 ~ a7 are joints of these parts. The bending angle of each finger calculated in the above process is

reflected directly in the length of each string. In this way, the angle of each joint of the marionette, corresponding to p1, p2, p3, p4, p5, p6, p7, and p8, is determined. Therefore, by bending each of the five fingers appropriately, a user can control the motion and gestures of a virtual CG marionette.

4.4 Background and CG Characters

We are planning a system that allows us to easily change scenes as well as characters, so we have developed various kinds of backgrounds and characters based on computer graphics. We are trying to develop an "Interactive Folktale System [8]" that offer users the ability to generate Japanese folktales as animation and to enjoy the interactions with creators of other characters in the system. Therefore, we have prepared various kinds of backgrounds and characters for our virtual marionette system. Figure 7 shows one of the marionette character in three different backgrounds.

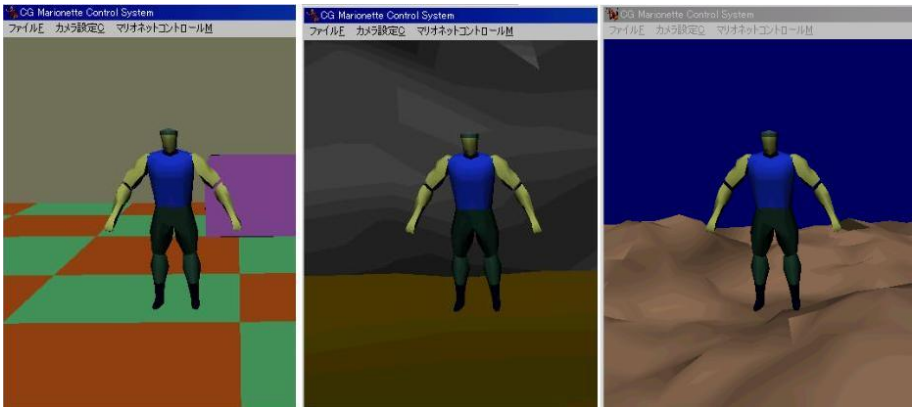


Fig. 7. Examples of virtual marionette characters

4.5 Networked Marionette Theater

The virtual marionette system we have developed as a first prototype toward the networked virtual marionette system consists of a hand-gesture recognition & control unit and an animation generation unit. This prototype system would work as a sub-system in the whole distributed system. These subsystems are connected using a network environment construction library called DirectPlay that is one of the DirectX libraries. In this case, instead of the client-server model, subsystems are connected based on the peer-to-peer model. In most of marionette plays the number of marionettes that appear in the play is less than ten. This means that this peer-to-peer network model would be enough to construct a networked marionette theater.

As a first step, we have constructed a networked system consisting of two sub-systems. The construction of the whole system is illustrated in Fig. 8. In each subsystem, the recognition results of another subsystem is shared. Furthermore, all of the CG characters and backgrounds are shared among these subsystems. Using these

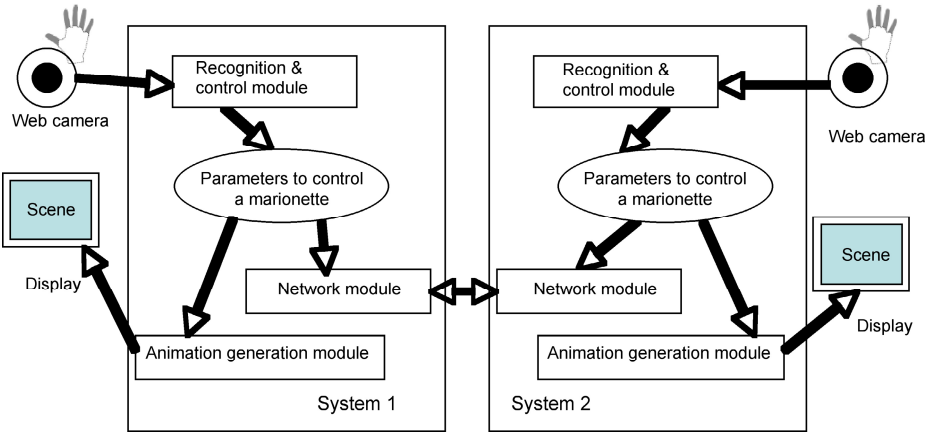


Fig. 8. Block diagram of two connected subsystems

recognition results as well as the CG characters and backgrounds, each subsystem can simultaneously create the same scene where multiple CG characters, each of which is controlled by its own subsystem, appear and behave in the same way.

5 Evaluation of the System

We have carried out an evaluation of a subsystem, which is the basis of the whole system and the instrument with which a user can control one virtual marionette character. We selected 20 students as subjects for this evaluation's tests. All of them know about marionette puppets but have never manipulated them. We asked them to manipulate both a real marionette puppet and a virtual CG marionette used in this system. After that we asked them several questions. The questions and the answers are summarized as follows.

(1) Is the movement of a virtual marionette "unique" compared with other CG characters?
Definitely Yes (4), Comparatively Yes (12), Neutral (4), Comparatively No (0), Definitely No (0)

(2) Is the movement of a virtual marionette "real"?
Definitely Yes (0), Comparatively Yes (1), Neutral (15), Comparatively No (4), Definitely No (0)

(3) Did you feel that your hand gestures were closely reflected in the movements of a virtual marionette?
Definitely Yes (0), Comparatively Yes (15), Neutral (3), Comparatively No (1), Definitely No (1)

From the first question, it is clear that 80% of the subjects said that there is some unique aspect in the movement of the virtual marionette. This means that the authors

succeeded in their intention to develop a system in which the particular movement of a marionette is regenerated. For the second question, the fact that most of the subjects answered "neutral" indicates that the meaning of "real" is somewhat difficult for them to associate with the marionette's movement. For the third question, 75% of the subjects answered that the marionette correctly moved according to their hand gestures. These results show that the recognition method introduced here works very well and gives people the feeling that they are directly manipulating the virtual marionette characters. Moreover, they again expressed the feeling that the system successfully reproduced the particular movement of a marionette.

6 Conclusions

In this paper, we proposed a system in which users can easily manipulate virtual marionette characters with their hand gestures. For the recognition of hand gestures, simple real-time hand-gesture recognition was realized by using histogram information of an extracted hand area. The recognition result is reflected in the movement of the marionette character by connecting each finger movement to a particular part of the virtual marionette by a virtual string. Furthermore, the concept of networked marionette theater was proposed in which several subsystems are connected by a network. Here, multiple users can perform theatrical marionette play by manipulating their own marionette characters. Finally, we carried out an evaluation test to assess the feasibility of a subsystem. By using twenty subjects and letting them manipulate both a physical marionette as well as a virtual one, we obtained evaluation results indicating that by using this virtual marionette system, even a non-expert of marionette manipulation can have the feeling of manipulating marionettes and thus can participate in a theatrical marionette performance.

For our further work, we need to improve the recognition accuracy of the hand-gesture recognition. Moreover, we need to develop adequate contents to refine the entire networked virtual marionette theater, and we also need to carry out an evaluation of the whole system by letting people use the system.

References

1. Keene, D.: No and Bunraku. Columbia University Press (1990)
2. http://www.lares.dti.ne.jp/bunraku/index.html
3. Currell, D.: Making and Manipulating Marionettes. The Crowood Press Ltd. (2004)
4. http://www.futuremovies.co.uk/review.asp?ID=319
5. Stam, J., Fiume, E.: Depicting Fire and Other Gaseous Phenomena Using Diffusion Process. In: Proceedings of SIGGRAPH 1995 (1995)
6. O' Brien, J.F., Hodgins, J.K.: Graphical modeling and animation of brittle fracture. In: Proceedings of SIGGRAPH 1999 (1999)
7. Courty, N.: Fast Crowd. In: ACM Siggraph/Eurographics SCA 2004 (2004)
8. Boulic, N., Thalmann, M., Thalmann, D.: A Global Human Walking Model With Real-Time Kinematic Personification. The Visual computer, 344–358 (1990)
9. Lee, J., Lee, K.H.: Precomputing avatar behavior from human motion data. Graphical Models 68(2), 158–174 (2004)

10. Ninomiya, D., Miyazaki, K., Nakatsu, R.: Study on the CG Marionette Control Based on the Hand Gesture Recognition. Annual Meeting of Game Society of Japan (in Japanese) (2006)
11. Ng, C.W.: Real-time gesture recognition system and application. Image and Vision Computing, 20 (2002)
12. Utsumi, A., Ohya, J., Nakatsu, R.: Multiple-camera-based Multiple-hand-gesture-tracking (in Japanese). Transaction of Information Processing Society of Japan 40(8), 3143–3154 (1999)
13. Miyazaki, K., Nagai, Y., Wama, T., Nakatsu, R.: Concept and Construction of an Interactive Folktale System. In: Ma, L., Rauterberg, M., Nakatsu, R. (eds.) ICEC 2007. LNCS, vol. 4740, pp. 162–170. Springer, Heidelberg (2007)

Tour into Virtual Environment in the Style of Pencil Drawing

Yang Zhao[1,2], Dang-en Xie[1], and Dan Xu[1]

[1] Yunnan University, Computer Science Department, Kunming 650091, China
[2] Yunnan Normal University, Computer Science Department, Kunming 650092, China
bootcool@163.com, xde820@gmail.com.cn, danxu@ynu.edu.cn

Abstract. Traditional virtual environment reconstructing methods have their own drawbacks, such as modeling complexity, higher time-consuming and much too realistic. Thus, they are difficult to be suitable for the real-time rendering. In this paper, we have described a system which can provide a simple pencil drawing scene model, the users can walk through in it freely, obtain enjoyable and real-time artistic experience. Firstly we design a new pencil texture generating method based on the pencil filter. This approach can conveniently generate the pencil drawing effect by convoluting the input image with the pencil filter. Secondly, we propose a modeling scheme for TIP based on the cubic panorama, which can not only overcome the disadvantage of fixed viewpoint in browsing panorama, but also can model easily and computes simply.

Keywords: pencil filter, pencil drawings, virtual environment, non-photorealistic rendering.

1 Introduction

Virtual environments simulate the visual experience of immersion in a 3D environment by rendering images of a computer model as seen from an observer viewpoint moving under interactive control by the user. It enables applications in education, computer-aided design, electronic commerce, and entertainment.

Current research in virtual environments uses traditional computer graphics theories to model and render virtual environments. The approach usually requires laborious modeling and expensive rendering hardware with special purpose. At the same time rendering quality and scene complexity are often limited because of real-time constraint. In addition, research in virtual environments has traditionally striven for photorealism, but for many applications there are advantages to non-photorealistic rendering (NPR). Firstly, artistic expression can often convey a specific mood difficult to imbue in a photorealistic scene. Secondly, through abstraction and careful elision of detail, NPR imagery can focus the viewer's attention on important information while downplaying extraneous or unimportant features [1].

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 628–635, 2008.

For these reasons, more and more researchers endeavor to depict virtual environments by NPR means. Allison W. Klein etc. have firstly described a very classic system for non-photorealistic rendering (NPR) of virtual environments [1].

Liviu Coconu etc. have also presents a NPR rendering pipeline that supports pen-and ink illustration for complex landscape scenes [2]. The feature of his work is that all NPR algorithms are integrated with photorealistic rendering, allowing seamless transition and combination between a variety of photorealistic and non-photorealistic drawing styles.

Thomas Luft etc. have presented algorithms that allow for real-time rendering of 3D-scenes with a watercolor painting appearance [3]. Their approach provides an appropriate simplification of the visual complexity, and imitates characteristic natural effects of watercolor. To enhance the efficient of the algorithm, they use image space processing methods rather than an exact simulation.

Hyunjun Lee etc. [4] have presented some real-time technique for rendering 3D meshes in the pencil drawing style. The shortage of the system is that the painting speed is limited by the amounts of polygon surfaces on the 3D model. Therefore, the computational complexity is very high when rendering the large scale and complex 3D scene.

In the research, we and Hyung W. Kang [5] have individually find that there are two disadvantages of these techniques mentioned above: (1) It has to build a coarse 3D model of the realistic 3D scene, so it is not fit to the real-time rendering of the complex scene. (2) Although many researchers have designed a various artistic style on NPR algorithm to render 3D virtual scenes, they ignored pencil drawing is an important form of pictorial representation and an effective way to convey lighting, directions and texture properties.

In this paper, by using a hybrid NPR/TIP approach which we designed previously [6], we present a new system for tour into a pencil drawing style virtual environment. By rendering the real photos or images into pencil drawing style, we can reconstruct the 3D virtual environment taking on artistic imago, at the same time we can easily transfer more concise and effective information to the user than any other artistic expressions practices.

2 System Overview

In this section we will simply introduce the basic architecture of our system. At a high level, our system proceeds in two steps. First, during preprocessing, the system has to do two important works (1) Reconstructs the foreground model and background model of the cubic panorama. (2) By using our image-based pencil drawing algorithm (This step is described in section 3), we can fast process the input 2D panorama images into pencil drawing styles. Second, during real-time rendering, the system has to do two important works (1) Re-renders the foreground model and background model in the style of pencil drawing (2) Interactively tour into the pencil drawing style cubic panorama based on the 3D model we created.

3 Image Based Pencil Drawing

Although the traditional pencil texture generating methods obtain good effect [7], they all have disadvantages of deficiency and time-consuming. Physical modeling is very complexity. LIC method needs to calculate the visualization vector field of the input image, and then convolute the pixel one by one, so it also costs much time. Generally, using the LIC method to generate a pencil drawing needs about 20 minutes (here the image size is 1024*768) [7, 8]. For these reasons, we proposed an image based techniques to directly process 2D images into pencil drawing styles. Our approach saves lots of time because pencil filters we designed are generated in advance and the convolution operation is more efficient than traditional methods. Figure 1 shows the framework of our pencil drawing algorithm.

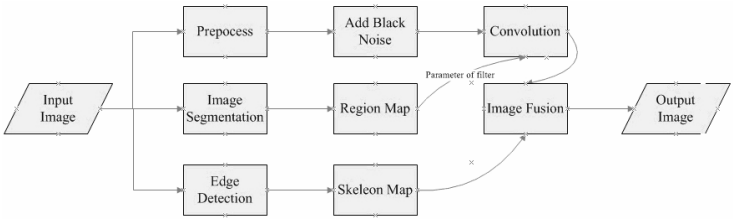


Fig. 1. The framework of the image based pencil drawing algorithm

3.1 Generating the Pencil Filter

By observing and analyzing the real pencil texture, we simply suppose that: (1) graphite marks present stochastic distribution according to the coarseness of papers; (2) graphite marks stretch along the stroke tracks; (3) graphite marks on perpendicular direction of a stroke present obviously black-white staggered distribution.

Considering the above supposes, we create a mathematic model for pencil filter. Assume that the stroke length is *len*, the stroke direction is θ and the stroke width is *2D*. If we know the stroke length and the stroke orientation, we can easily calculate the template size by $(\lceil len * \sin\theta \rceil \times \lceil len * \cos\theta \rceil)$.

The next problem is how to decide the value of each element in the pencil filter. As shown in Figure 2, firstly, calculate the distance *d* from each point *P* to the central axis *l* of a stroke. Then calculate the distance *r* from the point *P* to the center *O* of the stroke. The value of each element in the pencil filter lies on the relation between *d* and *D*, and also the relation between r and *len/2*.

Here, we take the upper right quarter of the template as an example (Figure 2(a)). Obviously, only three kinds of points are presented in the template. Points in the green area (e.g. P in Figure 2(b)) satisfy the conditions that *r* is less than *len/2* and *d* is less than *D*, so we choose *D-d* as their values. Points in the gray area (e.g. Q in Figure 3(c)) satisfy the condition that *d* is large than *D*, which means the stroke can't covered this area, so their values are set to be zero. Points in the blue area (e.g. R in Figure 3(d)), satisfy the conditions that r is large than *len/2* and *d* is less than *D*. Blue area is near to the stroke end, and the graphite marks is thin there, so the value in the

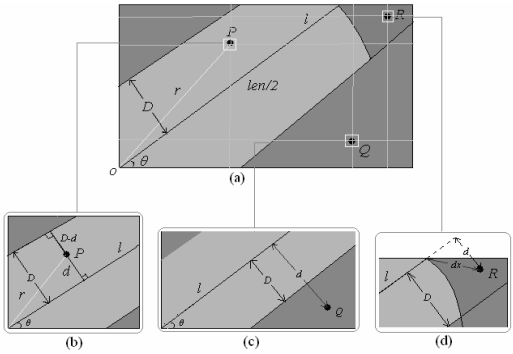


Fig. 2. Define a pencil filter

template is less than *D-d* and none zero. We calculate the distance *dx* from point *R* to the end of line l, and then we choose *D-dx* as the value of this area. In this way, it decreases the value of the stroke end effectively, and the decrement is in proportion to the distance *r*. When *D-dx* is less than zero, the value should be set to zero.

3.2 Generating the Black Noise Image

To make sure the pencil texture has stochastic distribution, we generate the black noise image from the reference image. Our method for generating the black noise image is similar to Mao's approach [7]. We use the tone of the input image to guide the distribution of the black noise. Let Iinput be the intensity of a pixel in the input image, P is a floating-point number generated with a pseudo-random function, and then the intensity Iinput of the corresponding pixel in the noise image is decided in the following way:

$$I_{noise} = \begin{cases} 255, & if \quad P \le T \\ 0, & otherwise \end{cases} \quad P \in [0.0,1.0], \quad T = k \cdot \left(\frac{I_{input}}{255} \right), \quad k \in (0.0,1.0] \tag{1}$$

K is a coefficient for controlling the density of the black noise. In this way we can ensure the pencil drawings have the stochastic distribution character, and also ensure the density of the black noise correspond to the intensity of the input image.

(a) (b) (c) (d)

Fig. 3. The contour lines maps with different value of μ

3.3 Extract the Contour Lines

It is also an important step to extrud the outlines in pencil drawing. Gradient operators are commonly used in digital image processing to extract edges of an image. Considering Kirsch operator has the bigger weighted factors, we prefer to choose it to extract the contour lines in this paper.

Here, μ is a coefficient for controlling the weight value in the filter. One can adjust the value of μ interactively. Figure 3 shows the different contour line maps with the different value of μ. Generally, if an image has more details, the value of should be smaller. A smaller can prevent the contour line conglutination. On the contrary, if an image has little details, a larger should be set in order to make sure the consistency of the contour lines.

3.4 Interactive Image Segmentation and Define the Stroke Orientation

In the process of pencil drawing, artists will choose strokes with different direction when painting a region. Therefore, our system needs to be done by the following two steps: (1) Segment the input image into different regions. (2) Set different stroke direction to the different regions. Traditional image segmentation techniques can not meet our requirements, so we have designed an interactive method for image segmentation. In our system, users can interactively segment the input image into different regions according to their creation needs, at the same time each of those regions will be then set to the corresponding stroke direction. Our approach is not only more flexible than Mao's approach [7], but also more in line with the actual process of pencil drawing. Figure 4(a) shows the input image. Figure 4(b) shows the result of the different segment regions, Figure 4(c) shows the stroke direction of the different regions.

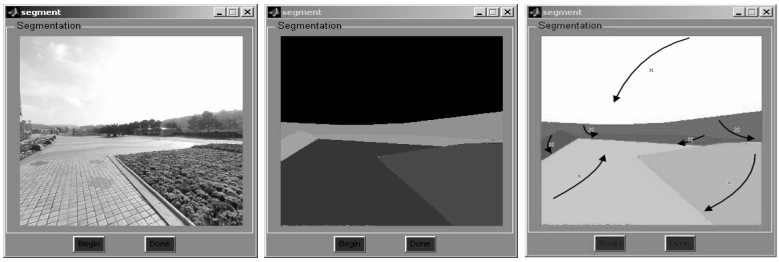


Fig. 4. Interactive Image segmentation and define the stroke orientation

4 Image Based Virtual Environments Modeling

To realize real-time interactive 3D pencil drawing virtual scene walkthrough, this section discuses the implementation technology of TIP [9] based on the vanishing line, and then describe our method of 3D cubic virtual scene walkthrough.

4.1 3D Scene Model Reconstruction Based on a Vanishing Line

According to the imagination and understanding to the input image, user can easily distinguish foreground from background. Thus the whole 3D scene model reconstructed mainly consists of background model and foreground model.

To background model, the input image is divided into two disjoint regions by the vanishing line. The region below the vanishing line in the image corresponds to ground plane, and that above the vanishing line corresponds to back plane. Since vanishing line contains all vanishing points formed by all parallel lines on a plane, back plane can be thought of as a plane of infinite distance and the points on it can be called the ideal points (Figure 5).To foreground model, a foreground object specified in the image is modeled as a 3D polygon called foreground model [9]. Suppose that the polygons stand perpendicular to the ground plane. The coordinates of its vertices are then computed by finding the intersection points between the ground plane and the ray starting from the viewpoint and passing through the corresponding vertices in the image plane. As proposed by H. Kang, a foreground object can have a hierarchical structure in a more complex environment. That is, another foreground object can be attached on a foreground object standing on the ground plane to form a hierarchical structure.

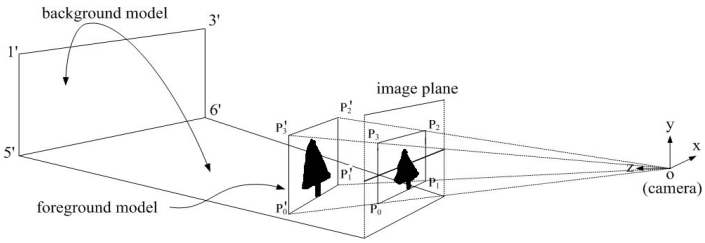


Fig. 5. Scene model based on a vanishing line

4.2 Tour into the Pencil Drawing Virtual Environments

For cubic panorama, there is no important content on the top and the bottom sides in general. Thus the system just simplifies to model the four sides using vanishing line based TIP techniques. Supposed that the center of cube (viewpoint) is positioned at the origin. Figure 6 illustrates the local model of one side of cube. We make local model for the four sides of cube respectively. Since the input image is panorama, the vanishing line on each of the four sides will have the same height above the bottom side. So the four local models will put together to form a global scene model. Then, the system just needs to project the top and the bottom sides of cubic panorama on the top and the bottom of the global model respectively. Finally, we get a close hexahedron scene model.

Together with the scene model, the images called background image and foreground mask are generated by segmenting the foreground objects from the input image. The background image is used as a texture, which is to be mapped onto the background model. Foreground mask is used to distinguish the exact portion of the

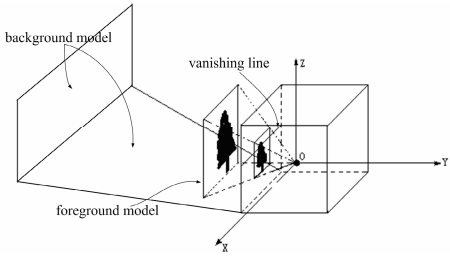


Fig. 6. The local model of one side of cube

foreground object from the background. After performing the above all steps, we can render the scene model by changing the parameters of camera.

5 Experiment Result

An image based pencil drawing system has been developed with the Matlab7.04. When the input image is specified, it can generate the pencil drawing image automatically. Users are allowed to specify parameters interactively. These parameters control the stroke orientation, length, density of the noise and the coefficients of the Kirsch operator. Compared with the LIC method, our method saves lots of time. We average use 10 seconds to process an image (here the image size is 1024*768) into pencil drawing style. The main reason is that we needn't to calculate the visualization vector field of the input image, and needn't to do the hundreds of iterations. We only need to do the convolution once for each pixel. In addition, we use Microsoft Visual C++ 6.0 and the OpenGL to develop our real-time interactive 3D

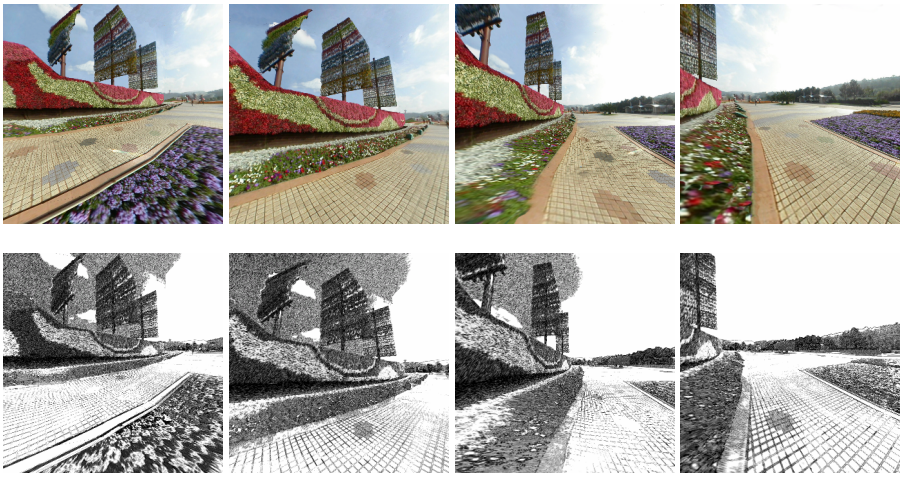


Fig. 7. Tour into virtual environment in the style of pencil drawing

virtual scene rendering system, the rendering speed of our rendering system can easily get to 25 fps. Through the combined use of the two systems we designed, users can tour into the virtual environment in the style of pencil drawing in real-time (Figure 7).

6 Conclusion

In this paper, we have described a system which can provide a simple pencil drawing scene model, the users can walk through in it freely, obtain enjoyable and real-time artistic experience. Our system is not only fit to the fast rendering of the small-scale scene, but also to that of the complex large-scale scene. Furthermore, we have also designed an image based system to directly process 2D images into pencil drawing style. According to their design requirements, users can make use of the system to render the input scene images into pencil drawing style. Through the combined use of the two systems we designed, this paper presents a relatively complete solution.

References

1. Klein, A.W., Li, W., Kazhdan, M.M., Corra, W.T., Finkelstein, A., Funkhouser, T.A.: Non-photorealistic virtual environments. In: ACM SIGGRAPH 2000, pp. 527–534 (2000)
2. Coconu, L., Deussen, O., Hege, H.-C.: Real-time pen-and-ink illustration of landscapes. In: NPAR 2006, pp. 27–35 (2006)
3. Luft, T., Deussen, O.: Real-Time Watercolor for Animation. J. Comput. Sci. Technol. 21(2), 159–165 (2006)
4. Lee, H., Kwon, S., Lee, S.: Real-time pencil rendering. In: NPAR 2006, pp. 37–45 (2006)
5. Kang, H.: Nonphotorealistic Virtual Environment Navigation from Images. International Journal of Image and Graphics 5(2), 1–13 (2005)
6. Zhang, Y.-P., Zhao, Y., Shi, J., Xu, D.: Digitization of Culture Heritage Based on Tour into the Picture. In: Edutainment 2006, pp. 1312–1320 (2006)
7. Mao, X., Nagasaka, Y., Imamiya, A.: Automatic generation of pencil drawing from 2D images using line integral convolution. In: CAD/Graphics 2001, pp. 240–248 (2001)
8. Cabral, B., Leedom, C.: Imaging Vector Field Using Line Integral Convolution. In: SIGGRAPH 1993 conference proceeding, pp. 263–270 (1993)
9. Kang, H., Pyo, S.Y., Anjyo, K., Shin, S.Y.: Tour into the Picture using a Vanishing Line and its Extension to Panoramic Images. Computer Graphics Forum 20(3), 132–141 (2001)

Research and Implementation of Hybrid Tracking Techniques in Augmented Museum Tour System

Hong Su, Bo Kang, and Xiaocheng Tang

Mobile Computing Center, College of Automation Engineering,
University of Electronic Science and Technology of China, Chengdu, China
suhongasr@gmail.com, kangbo@uestc.edu.cn, christile@126.com

Abstract. Augmented Museum Tour (AMT) system aims at providing a vivid and interactive approach by using Augmented Reality Techniques (ART) for museum wandering, but few effective tracking methods were taken into research previously. In this paper, we try to apply a hybrid tracking approach which integrates inertial (6-DOF) and vision-based technologies to promote the tracking performance for Augmented Museum Tour system, followed by analyzing the tracking problems and requirements for it. An AR experimental system based on VisTracker, a Vision-Inertial Self tracker is also designed. Experimental results and analysis demonstrate the system's effectiveness and its further application prospect in Augmented Museum Tour system.

Keywords: Augmented Reality; Hybrid Tracking; VisTracker; Augmented Museum Tour system.

1 Introduction

Augmented Reality (AR) is a new research field developed on the basis of Virtual Reality (VR), which superimposes computer generated virtual information onto the surrounding real environments to augment user's perception in real-time and interactively[1]. Augmented Reality systems have a promising application prospect[2] in the field of Equipment Maintenance, Entertainment, Medical treatment and E-learning.

By applying modern AR techniques into conventional museum tour[3-5], named as Augmented Museum Tour system (AMT), it becomes more effective and interactive for visitors, helps them absorb profound information and experience a special multi-channel interactive way, including panorama visual effect and audio approaches. Although Augmented Museum Tour system may benefit visitors a lot, just like all AR prototype systems, it also needs to solve tracking problems first to achieve a better performance.

Tracking is one of important research fields in Augmented Reality (AR), and also in Augmented Museum Tour (AMT) system; it aims at solving accurate registration problems between virtual information and real environment and building a stable and effective AR application system. Former researches on tracking techniques in Augmented Museum Tour system mainly focus on vision-based technology. Noboru

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 636–643, 2008.

Koshizuka[4] proposed an architect to navigate in Museum Environments to measure the distance between visitor and material by using vision-based tracking technology and Fotis Liarokapis[5] also provided a prototype for museum wandering by recognizing markers with CCD camera. Although it works well in an experimental environment, vision based tracking has shortcomings in limited tracking range and high calculation cost, a more effective hybrid tracking approach will be needed upon the tracking requirements of Augmented Museum Tour system.

2 Tracking Requirements for Augmented Museum Tour System

Augmented Museum Tour (AMT) system aims at providing a collaborative and assistant platform for visitors. Based on the practical application needs and sensitivity of human eyes, there will be high standards required on the accuracy of registration and realistic rendering effect of virtual information. The requirements[6] for better tracking performance of AMT system are as follows:

Accuracy: the accuracy of position measurement should be limited to 0.1 centimeters, and the orientation measurement to 0.1 degrees;

Real-time: the general response time and delay time should be as short as possible. For ideal AR systems, the delay time should be shorter than 1 millisecond, and for general AR applications, it should be in 10 milliseconds.

Tracking range: AR system always works in real 3D environments. To be used in a mobile computing environment, large-scale tracking areas are required for users to interact with real environment and virtual information under various lighting conditions.

Robust: The ability to maintain an effective tracking performance under surrounding changing environments, just like occlusions or noise interference.

Besides, size, weight, and cooling ability factors should also be taken into consideration to keep a better mobile computing performance.

Recent researches demonstrate that there are three tracking techniques mainly used in AR system, including magnet-based tracking, vision-based tracking and inertial-based tracking techniques. For the complexity of practical use in AR, there are no single tracking techniques which could satisfy all the tracking requirements described above. Hybrid tracking attempts to compensate for the shortcomings of each technology by using multiple measurements, which will be the main research approach to achieve better and effective tracking results in Augmented Museum Tour system.

3 Vision-Inertial Based Hybrid Tracking

Because of deficiencies of vision-based tracking in stability and real-time performance, Suya You[7] introduces inertial tracking into vision-based AR systems, fusing 3DOF inertial orientation data with vision features to stabilize performance and furthermore correct inertial drift.

In this paper, a 6DOF inertial sensor is applied as the hybrid tracker into our system. The hybrid tracking process consists of three procedures. First, the initial position of the tracker is computed by visual recognition; Second, inertial tracking data will be

used for predicting the dynamic position of the tracker; Third, visual assisted correction is implemented to ameliorate inertial drift and compute the accurate position and orientation information of the tracker.

3.1 Initial Positioning with Vision Tracking

Initial positioning will be first performed by recognition with vision tracker after the initialization of the system. Just like every visual recognition algorithm, markers in image sequences are first sampled for pre-processing, four to five best sampled markers are detected and extracted for recognizing, and the position information of these markers are looked-up in the pre-stored map, which will be used for computing the initial position and orientation information with pose recovery algorithm.

3.2 Dynamic Positioning with Inertial Tracking

Dynamic positions will be predicted according to the inertial tracker. Provided by inertial tracker, acceleration vectors and rotation rates are combined together to predict the approximate area of the tracker in reference coordinator by Data Fusing Processor.

3.3 Visual Assisted Correction

Visual data are also used to ameliorate inertial drift by using Kalman filter and acquire the accurate position information in the approximate area provided by inertial tracking step.

3.4 Coordinate System

There are four principal coordinate systems in Augmented Museum Tour system, as illustrated in Figure 1: the world coordinate system $W(x_w, y_w, z_w)$, the camera coordinate system $C(x_c, y_c, z_c)$, the inertial coordinate system $I(x_i, y_i, z_i)$, and the 2D image coordinate system $U(x_u, y_u)$.

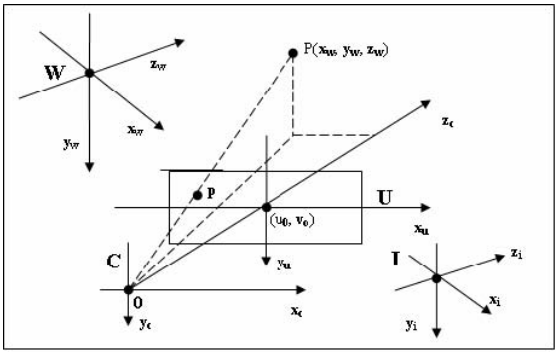


Fig. 1. Coordinate systems of the hybrid tracking

A pinhole camera models the coordinate transforming process. The transformation from a point $P(x_w, y_w, z_w)$ in world coordinate system to 2D image coordinate system is:

$$W: \rightarrow U: \quad Z_c \begin{bmatrix} x_u \\ y_u \\ 1 \end{bmatrix} = K \begin{bmatrix} R_i & T_i \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \tag{1}$$

where the R_i is a 3*3 rotation matrix, calculated with orientation data from gyroscopes and the T_i is a 3*1 translation vector, calculated with position data from accelerometers, characterizing the current orientation and position of camera in world coordinate system.

where the matrix K:

$$K = \begin{bmatrix} f/\Delta x & 0 & u_0 \\ 0 & f/\Delta y & v_0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2}$$

contains the intrinsic parameters of the camera, f is the focal length of the camera, Δx and Δy are the physical size of horizontal and vertical axis in 2D image plane, and $(u0, v0)$) is the intersection point of camera-center axis with 2D image plane. The intrinsic parameters are calculated offline.

4 VisTracker Based Hybrid Approach in Augmented Museum Tour Experimental System

Our experimental system aims at simulating the design and implementation of Augmented Museum Tour system, taking the showcase in the lab as the augmented object and superimposing virtual 3D information onto it to add more profound information. While engaging in the simulation, hybrid tracker is used to track user's position and its line of sight and measure the distance between user and the showcase in real time.

The virtual 3D information will be rendered at the specified point with corresponding orientation at the moment the user is looking toward the direction of the showcase.

4.1 System Platform

IS-1200 VisTracker[8] is used as our hybrid tracker. VisTracker designed by InterSense.Inc in 2003 is a Vision-inertial self Tracker, which could achieve accurate and real-time tracking performance in a large and various lighting environment. Its better performance can satisfy the tracking requirements for our application needs in Augmented Museum Tour system.

VisTracker, a 6DOF hybrid tracker, has two standard units[9]. One is hybrid sensor unit, used for recognizing markers and acquiring inertial information during moving process; the other is Data Fusion processor, used for computing the position and orientation information of the tracker in world coordinate system.

Our system takes ARToolkit[10] released by HIT Lab from University of Washington as our software developing platform, OpenGL 3D Graphics API as virtual information rendering tool; Hardware platform is made up of three parts: VisTracker, CCD camera and a computer workstation. VisTracker and CCD camera are fixed together onto a bracket with their centers calibrated in one point and a 90° offset in the pitch angle.

4.2 System Design

System design introduces the way we had applied to implement our system.

4.2.1 Setup of World Coordinate and Disposition of Markers

The position and orientation information of VisTracker are obtained by computing the coordinate position of markers. In our system, a world coordinate system is first setup in user's working space with its origin on the ceiling and Z-axis pointing down the ceiling plane.

Feature markers are disposed to the XY plane on the ceiling plane. There are seven fixed markers in our system, named from ID100 to ID400, and their positions in world coordinate system are also carefully measured in Figure 2:

4.2.2 Position Measurements for Showcase

The position of showcase in the lab is obtained by carefully geometric measurement, set as $P(x_w, y_w, z_w)$ in world coordinate system.

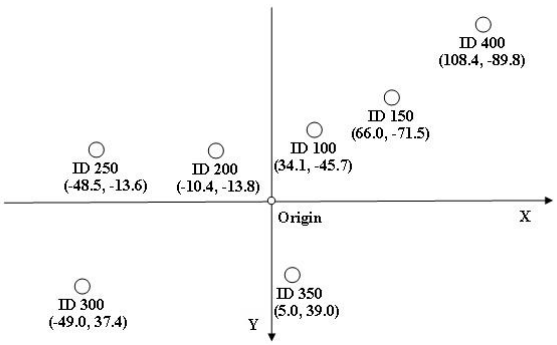


Fig. 2. World coordinate system. The metric is in centimeter, and *z-axis* forwards inside the paper.

4.2.3 Acquiring Position and Orientation Information

The position and orientation information of VisTracker in world coordinate system are acquired directly by using SDK provided by VisTracker Develop Kit, set as

$V(x_p, y_p, z_p)$ and $V(x_o, y_o, z_o)$ respectively and according to it, we can get the camera position $C(x_p, y_p, z_p)$ and orientation information $C(x_o, y_o, z_o)$ under world coordinate system automatically.

4.2.4 Calculating Extrinsic Parameters

Equation (1) demonstrates that the extrinsic parameters $[R_i, T_j]$ of the camera should be calculated first to transform the world coordinate into 2D image coordinate, where the R_i is a 3*3 rotation matrix, calculated with orientation data from gyroscopes and the T_i is a 3*1 translation vector, calculated with position data from accelerometers. The equations are as follows:

Rotation matrix:

$$R_i = \begin{bmatrix} \cos(x_0)*\cos(z_0)+\sin(x_0)*\sin(y_0)*\sin(z_0) & -\cos(x_0)*\sin(z_0)+\sin(x_0)*\sin(y_0)*\cos(z_0) & \sin(x_0)*\cos(y_0) \\ \cos(y_0)*\sin(z_0) & \cos(z_0)*\cos(y_0) & -\sin(y_0) \\ -\sin(x_0)*\cos(z_0)+\sin(y_0)*\cos(x_0)*\sin(z_0) & \sin(z_0)*\sin(x_0)+\sin(y_0)*\cos(x_0)*\cos(z_0) & \cos(x_0)*\cos(y_0) \end{bmatrix} \quad (3)$$

Translation vector:

$$[T_i] = \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix}. \quad (4)$$

4.2.5 Rendering Virtual 3D Information

After the extrinsic parameters calculation process, OpenGL 3D Graphics API is then used to render virtual augmented information into video sequences at the specified point with corresponding orientation on the software developing platform, then user could experience such emerging effects according to the computer screen.

4.3 Experimental Results and Analysis

In our experimental system, user implements Augmented Museum Tour system by holding the VisTracker under the calibrated space fixed with markers, just like the process visitors wandering through the museum.

Figure 3 demonstrates a working environment and Figure 4 shows the experimental results where only teapot is the computer generated virtual object acting as assistant information for the showcase. Computer will render virtual object automatically while the user is looking toward its direction.

Experimental results show that vision-based inertial hybrid tracking achieves a good performance under large and various lighting conditions, with a shorter time delay in both tracking and rendering process, and performs a high tracking accuracy, with position to 0.1 centimeters and orientation in 0.1°. However, VisTracker is not suitable for long time use because of its higher thermal power.

Fig. 3. Working environment

Fig. 4. Experimental results

5 Conclusion

We tried to apply a hybrid tracking approach which integrates inertial (6-DOF) and vision-based technologies to promote the tracking performance for Augmented Museum Tour system, followed by analyzing the tracking problems and requirements for it. An AR experimental system based on VisTracker, a Vision-Inertial Self tracker is also designed. Experimental results and analysis demonstrate the system's effectiveness and its further application prospect in Augmented Museum Tour system.

References

[1] Azuma, R.: A Survey of Augmented Reality. J. Presence: Teleoperators and Virtual Environments 6(4), 355–385 (1997)
[2] Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre, B.: Recent Advances in Augmented Reality. J. IEEE Computer Graphics and Applications, 34–47 (2001)
[3] Vlahakis, V., Ioannidis, N., Karigiannis, J.: ARCHEOGUIDE: Challenges and Solutions of a Personalized Augmented Reality Guide for Archaeological sites. J. IEEE Computer Graphics and Applications 22(5), 52–60 (2002)
[4] Koshizuka, N.: Museum Navigation System using Augmented Reality Technologies, `http://www.um.u-tokyo.ac.jp/publish_db/2000dm2k/english/01/01-16.html`
[5] Liarokapis, F., White, M.: Augmented Reality Techniques for Museum Environments. J. The Mediterranean Journal of Computers and Networks 1(2), 90–96 (2005)
[6] Kang, B.: Tracking Technology for Augmented Reality. Computer Measurement and Control 14(11), 1431–1434 (2006)
[7] You, S., Neumann, U., Azuma, R.: Hybrid inertial and vision tracking for augmented reality registration. In: Proc. of IEEE Virtual Reality 1999, pp. 260–267 (1999)

[8] Foxlin, E., Naimark, L.: VIS-Tracker: A Wearable Vision-Inertial Self-Tracker. In: Proc. of IEEE Virtual Reality 2003 (VR 2003), pp. 199–206 (2003)
[9] Naimark, L., Foxlin, E.: Circular Data Matrix Fiducial System and Robust Image Processing for a Wearable Vision-Inertial Self-Tracker. In: IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2002), pp. 27–36 (2002)
[10] Kato, H., Billinghurst, M., Blanding, R.: ARToolkit PC version 2.11., http://www.hitl.washington.edu/artoolkit/

Terrain Synthesis Based on Microscopic Terrain Feature

Shih-Chun Tu, Chun-Yen Huang, and Wen-Kai Tai

Department of Computer Science and Information Engineering,
National Dong Hwa University, Taiwan, Republic of China.
1, Sec. 2, Da Hsueh Rd., Shou-Feng, Hualien, 974
Taiwan, Republic of China
tusjtu@ms01.dahan.edu.tw,
NightSun@game.csie.ndhu.edu.tw,
wktai@mail.ndhu.edu.tw

Abstract. In this paper, we use real terrain elevation data and the microscopic terrain feature unit to synthesize the macroscopic terrain flexibly and effectively. An interactive system provides users a convenient and intuitive interface to profile microscopic terrain features using terrain primitives. A number of terrain primitives, geometric objects such as prism-like, tetrahedron-like, and cone-like objects regarded as the concept representation object of microscopic terrain features, can be transformed to construct the terrain profile. Then, these terrain primitives are replaced by matching terrain units searched from real terrain elevation database to seamlessly synthesize the macroscopic terrain, the landform. As experimental results show, the resultant efficient synthesized terrains are realistic, fitting user's intuition.

Keywords: Terrain Modeling, Terrain Synthesis, Terrain Primitive.

1 Introduction

Terrain synthesis is essential to the construction of realistic outdoor virtual environments. Many applications such as pilot training, scenery browser, game, etc. need terrain models. Upon terrain synthesis, fractal terrain approaches and physical erosion approaches were two main techniques in the past few years. Fractal terrain approaches based on the Brownian motion lack realism partly because the statistical character of the surface is the same everywhere, i.e. the surface has no global erosion features inherently. Physical erosion approaches simulate fluvial, thermal, and diffuse erosion processes to create global stream/valley networks, talus slopes, and rounding of terrain features, so they are computation intensive due to the complex physical model.

In this paper, we use several microscopic terrain features to construct a macroscopic terrain. Microscopic terrain features [1] are hill, mountain, plain, tableland and plateau. They are classified by elevation, relative altitude and gradient. We use several terrain units to construct a microscopic terrain feature. The macroscopic terrain represents a scene namely, a landform, being constructed by one or

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 644–655, 2008.

more microscopic terrain features. We provide users an interactive environment to profile microscopic terrain features using terrain primitives. These user specified terrain primitives are replaced by real terrain units from terrain elevation database. External and internal matching criteria used by proposed matching algorithm are employed to evaluate real terrain units according to attributes of terrain primitive. The best matching terrain units then substitutes for the corresponding terrain primitives. Finally, stitching is carried on, if necessary, adjacent terrain units for having optimally synthesized macroscopic terrain.

The proposed interactive environment provides an effective and convenient way such that user can intuitively specify a microscopic terrain profile using terrain primitives. As compared to commercial tools such as Bryce, WorldBuilder, etc. from which users synthesize terrain by editing height map directly. The synthesized terrain of our approach is more realistic and reasonable than that of fractal terrain approaches and physical erosion approaches, because all terrain units are from real world. Also, our approach is more efficient because ours does not proceed simulation but searching for the best matching terrain unit.

The paper is based on Chiang [19], the comparisons and improvements between the paper and Chiang [19] are listed below:

1. In the paper, we use the geometry and topology similarity to simplify the matching procedure and exploit an alpha value to adjust the candidate set of the terrain units that have the most geometry and/or topology similarity such that the new method is more flexible and effective.
2. In Chiang [19], the matching method often obtains a strictly small number of candidate terrain units and results in the stitching process performing not well.
3. In the paper, we construct a graph to find the shortest path between all candidate terrain unit sets such that we can have a global matching result rather than just locally deciding the best matched unit in Chiang [19]. Therefore, we have synthetic terrains with better visual effect as you can see results shown in the experimental section.
4. In short, the paper is more flexible and effective.

The rest of the paper is organized as follows. Section 2 reviews previous work on terrain generation. Section 3 describes our approach. Section 4 demonstrates the experimental results of our terrain synthesis method, and finally we give some conclusions and future works in section 5.

2 Related Works

In the past few years fractal terrain approaches, [2],[3],[4],[5],[6],[8],[9],[10], [13],[15], and physical erosion approaches, [11],[12], were two main approaches in terrain synthesis.

Lewis [4] generalized the stochastic subdivision, constructed by Fournier, Fussell, and Carpenter [2], based on the random process and estimation theories to synthesize a noise with prescribed autocorrelation and spectrum functions.

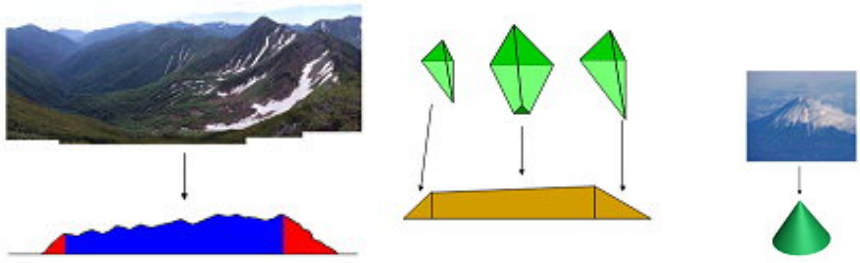
Lewis produced artifact-free noises with a variety of spectra, and the gray levels in synthesized textures are interpreted as heights to obtain several distinct types of synthetic terrains. The Poisson faulting process proposed by Mandelbrot [5] is a sum of randomly placed step functions with random heights. It produced a Brownian surface. While it is suitable for use on spheres for creation of planets, yet with $O(n^3)$ time complexity. The variation of stochastic subdivision described by Miller [9] used a higher quality interpolation for the point estimation that alleviated the problem of creases. Musgrave [10] generated fractal terrains by terrain generation and erosion simulation. In terrain generation phase, Musgrave used summing band-limited noises which refer to as noise synthesis to generate fractal terrain. In erosion simulation phase, Musgrave subdivided the erosive processes into two categories: hydraulic erosion and thermal weathering. The approach improves on previous fractal terrain synthesis and rendering techniques and the result looks very realistic. Nagashima [11] used fractional Brownian motion (fBm) obtained mainly from Poisson faulting by Mandelbrot [5] and Peitgen and Saupe [13], and Fourier filtering by Peitgen and Saupe [13] for creating fractal terrain models so that the approach could produce pictures of realistic terrains. However, if the fBm and related fractal methods such as Mastin [8] are used independently, it will produce unnatural terrains.

Physical erosion approaches are based on physical erosion theory. Nagashima [11] proposed a simulation approach based on alluvial deposition laws, hydraulic and thermal watering erosion. This approach shows amazing results. Roudier [12] proposed a terrain simulation model for erosion and deposition using geological parameters and laws from simplified geomorphological theories. However, the approach only considered the processes that depend on running and infiltrated water, and the timing costs a lot.

There are other methods which do not use fractal approach or physical simulation to generate terrains. Marshall [7] proposes a procedural model to generate scenes automatically. The procedural model uses fundamental data parameters to create object, and the data parameters are always trivial for user. So this method is useful even though its results are decided by the complete of procedural model.

3 Approach

In this section, we specify the proposed approach for synthesizing macroscopic terrains. First, user profiles a microscopic terrain in our interactive environment using terrain primitives. Second, for specified terrain primitives, the matching procedures, external and internal procedure, are applied to search for the best matching terrain units from the terrain unit database. Third, the best matching terrain units constitute for the corresponding terrain primitive. Finally, to have the best smoothness and visual effect of the synthetic terrain, the minimum cut approach is exploited to find an optimal seam such that adjacent terrain units are well stitched.



(a) (b) (c)

Fig. 1. (a) A mountain terrain can be divided into two boundary zones (red) and one intermediate zone (blue). (b) Two tetrahedron-like geometry primitives and one prism-like geometry primitive are suitable to conceptualize the mountain terrain. (c) The cone-like geometry primitive can be used to conceptualize this type of mountain well.

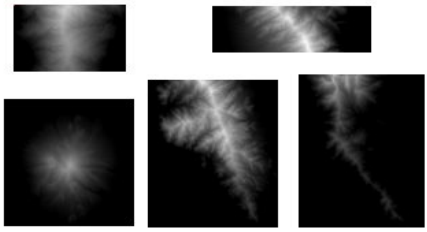
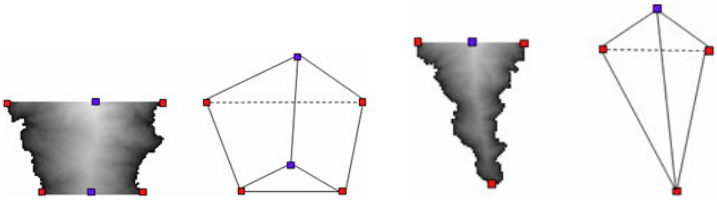


Fig. 2. Sample terrain units



(a) (b)

Fig. 3. Two sample terrain units. (a) The prism-like terrain primitive is appropriate to conceptualize the terrain unit at left side. (b) The tetrahedron-like terrain primitive well profile to its terrain unit.

3.1 Terrain Primitive

A terrain primitive is a concept representation object being used to profile the microscopic terrain feature. Geometrical primitives such as prism-like,

tetrahedron-like and cone-like object are suitable to be terrain primitives. Generally more than one terrain primitive are necessary to completely compose a microscopic terrain feature. In our interactive system, we have users flexibly construct a target microscopic terrain feature using drag-and-drop mechanism, allowing translation and scaling transformation on specified terrain primitives.

While constructing, for instance, user can conceptualize a mountain range as two boundary zones and one intermediate zone as shown in Figure 1(a). The tetrahedron-like terrain primitive can represent the boundary zone, and the prism-like terrain primitive fits the intermediate zone as shown in Figure 1(b). Figure 1(c) shows another type of mountain which we use the terrain primitive, cone-like object, to conceptualize it. Note that variants of a terrain primitive can be obtained by applying geometric transforms to the primitive. Therefore, to profile most possible forms of mountains, those terrain primitives are sufficient and effective.

3.2 Terrain Unit

The terrain unit database contains as wide range of all types of microscopic terrain features, like hill, mountain, plain, tableland and plateau, as possible such that the matched terrain unit to user specified terrain primitive is as close as possible. All terrain units are manually segmented from terrain elevation map of real world according to rules as follows:

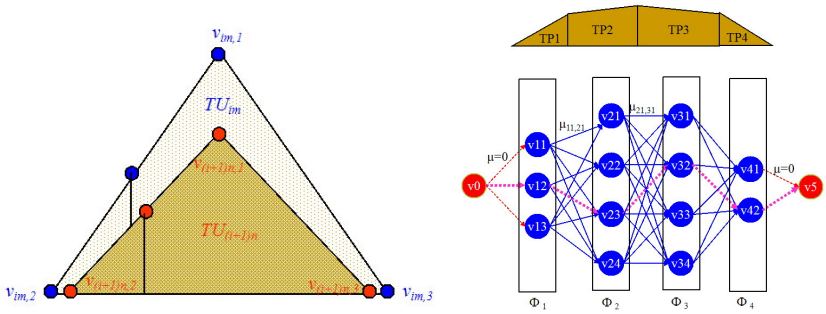
1. The height variation of each scanline in the terrain unit displays higher elevation near the center and lower elevation around the boundary.
2. Each terrain unit contains only one main mountain ridge. This requirement makes user specify the terrain primitive easier, and so does the matching process.

Sample terrain units are shown in Figure 2. All segmented terrain units are oriented to align each other and normalized in height and width length. Besides, each terrain unit is conceptualized to correspond a terrain primitive using its top and bottom cross sections and main range. Two sample terrain units with corresponding terrain unit primitive are shown in Figure 3.

3.3 Matching Algorithm

After user specifies terrain primitives in our system, the best matching terrain unit is going to be searched to substitute for the terrain primitive. There are two phases. First, the external matching phase finds out candidate terrain units that have the most geometry and topology similarity to the specified terrain primitive. The internal matching phase then determines the best matching terrain unit from each candidate set.

External Matching. External matching measures the geometry and topology similarity. To measure the geometry similarity, we compare the vertices of



(a) (b)

Fig. 4. (a) Two connected cross sections of adjacent terrain units are aligned with the peak position. (b) The microscopic terrain is constructed by four terrain primitives and a graph $G = (V, E)$ is constructed to find the shortest path (pink) from $v0$ to $v5$.

specified terrain primitive with that of terrain unit primitive in the database using

$$G_{i,j} = \sum_{k=1}^{4or6} |TP_i(v_k) - TUP_j(v_k)|$$

where $G_{i,j}$ is a geometrical distance from the terrain primitive (TP) i to the corresponding primitive of a terrain unit (TUP) j, and $TP_i(v_k)$ and $TUP_j(v_k)$ indicate the kth vertex of TP and TUP respectively. To measure the topology similarity, the K-DOP (k Discrete Oriented Polytope) [14] of terrain primitive and terrain unit (TU) are measured by

$$T_{i,j} = \sum_{k=1}^{26} |TP_i(P_k) - TU_j(P_k)|$$

where $T_{i,j}$ is a topological distance from TP_i to TU_j, and $TP_i(P_k)$ and $TU_j(P_k)$ indicate the corresponding kth plane of TP_i to TU_j respectively. Given a threshold ϵ, a candidate set of terrain units for a user specified terrain primitive is

$$\Phi_i = \{TU_j | (\alpha g_{i,j} + (1-\alpha) t_{i,j}) \leq \epsilon, 0 \leq \alpha \leq 1\}$$

where $g_{i,j}$ and $t_{i,j}$ are the normalized $G_{i,j}$ and $T_{i,j}$ respectively. The geometry distance metric, $g_{i,j}$, is used to evaluate the cross-section and mountain ridge similarity and the topology distance metric, $t_{i,j}$, measures terrain contour (shape) similarity.

Internal Matching. When obtaining the candidate set Φ_i, we determine the best matching terrain unit from Φ_i, which has the lowest cost of smoothness.

Since the microscopic terrain is composed of real terrain units, say I, the joint of two adjacent terrain units matters the visual smoothness while rendering the synthesized terrain. A smoothness cost measures the smoothness at joint as follows:

$$u_{i,i+1} = \sum_{k=1}^{3} |v_{i,k} - v_{i+1,k}| + MiddleSmotthCost(v_{i,1}, v_{i,2}, v_{i+1,1}, v_{i+1,2}, d, w) + MiddleSmotthCost(v_{i,1}, v_{i,3}, v_{i+1,1}, v_{i+1,3}, d, w)$$

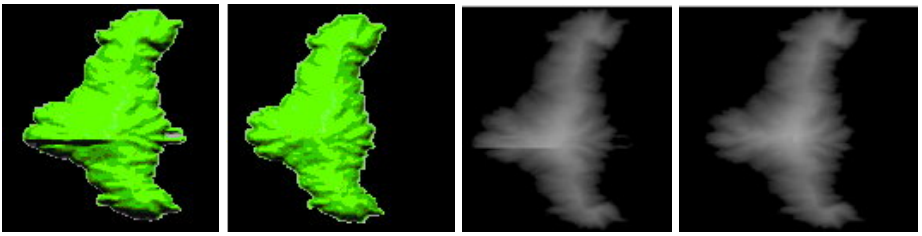
In the first term, we concern the degree of elevation differences at three major extreme vertices of two adjacent terrain units, say two end vertices and the vertex with highest elevation at the adjacent cross section of a terrain unit as shown in Figure 4(a). We further consider elevation variations between two pairs of major vertices in the second term, *MiddleSmoothCost* function. Parameter d controls the fineness of smoothness cost measurement between two endpoints, $v_{x,1}$ and $v_{x,2}$ or $v_{x,1}$ and $v_{x,3}$ where x is i or $(i+1)$, namely the depth of recursion. The last parameter, w, weights the importance for each measurement. In our experiment, the weighting w is assigned to 0.5, and the depth d is assigned to 10, i.e. 512 pairs of vertex will be measured. The pseudo code of the function is shown as following.

```
procedure MiddleSmoothCost(v1, v2, v3, v4, d, w)
   {recursion stops when the weight value of w is
    less than a given threshold t.}
   if (w < t) return (0);
   {recursion stops when the distance between two endpoints is
    less then a given theshold c.}
   {v1.x represents x coordinate of v1.}
   if (|v1.x - v2.x| < c) or (|v3.x - v4.x| < c) return (0);
   if (d = 0) return (0);
   {midpoint1 and midpoint2 are middle point of v1, v2 and
    v3, v4 respectively.}
   midpoint1.x=(v1.x+ v2.x) / 2;
   midpoint2.x=(v3.x+ v4.x) / 2;
   {v.height represents height of v.}
   return w * | midpoint1.height - midpoint2.height | +
          MiddleSmoothCost(midpoint1,v2,midpoint2,v4,d-1,w*w) +
          MiddleSmoothCost(midpoint1,v1,midpoint2,v3,d-1,w*w);
end.
```

The smoothness cost is concurrently evaluated for all candidate terrain units in each Φ_i with respect to a specified microscopic terrain feature. A graph $G =< V, E >$ is constructed, where V is the union set of all elements in Φ_i for all $i \in I$ with two extra vertices, say source vertex and target vertex, and for all

pairs of candidate terrain units in adjacent candidate set Φ_i and Φ_{i+1}, there is an edge which is associated with the measured smoothness cost. In Figure 4(b), for example, one constructs a microscopic terrain with four terrain primitives, TP_1, TP_2, TP_3, and TP_4. Let Φ_i be the candidate set of terrain units of $TP_i, i = 1, 2, 3, 4$ respectively. Two extra vertices, a source vertex ($v0$) and a target vertex ($v5$), are added in the G, where the smoothness cost is zero for all edges from $v0$ to all candidate terrain units in Φ_1 and from all candidate terrain units in Φ_4 to $v5$. The shortest path from $v0$ to $v5$ is found by Dijkstra shortest path algorithm [16]. The vertices $v12$, $v23$, $v32$ and $v42$ on the path are the best matching terrain unit of TP_1, TP_2, TP_3 and TP_4 respectively. In a short, we employ the internal matching method to promise one optimal terrain unit set from which the microscopic terrain feature demonstrates the best visual quality and hopefully best fit to users intuition

Notice that users are allowed to assign a terrain unit from Φ_i as the best matching unit of TP_i if they prefer some specific terrain features. If so, the shortest path algorithm just applies to a graph which has some cut vertices (user assigned terrain unit). The dark side of this assignment is that we may have a shortest path where the terrain appearance might reveal some discontinuities at joint boundaries.



(a) (b) (c) (d)

Fig. 5. (a) Terrain units join together without stitching. (b) Terrain units join together using the minimal cut approach for stitching. (c) and (d) are heightfield maps of (a) and (b) respectively.

3.4 Terrain Unit Placement and Joint

On obtaining the best matching set of terrain units for the specified microscopic terrain feature, they must be placed on the constructing terrain for synthesis. In addition to rotation and translation transformations applied for aligning with the orientation of the terrain primitive, scaling however must be performed occasionally to the terrain unit because its dimensions might not be the same with the terrain primitive. Scaling does not change the terrain appearance much because the matching terrain unit is the closest one by the similarity measurement rules.

Simply substituting for specified terrain primitives may not make the synthetic terrain look smoothness. Even for best matching terrain units, stitching for two adjacent ones is still required. Let two adjacent terrain units be placed partly overlapped. Our goal of stitching is to find the minimum cut at the overlap region such that adjacent terrain units join together as seamless as possible. Inspired by Kwatra et al. [17], we can simply employ the minimum cut [18] to find a good seam, and then adjacent terrain units are stitched together along this seam. Furthermore, to obtain the optimal seam, we first vary the overlay region by iteratively offsetting one terrain unit within a range, and compute the sum-of-height-differences cost [17] for the offset overlay region at each iteration, and finally find the minimum cost at the offset overlap region with the lowest sum-of-height-differences cost.

Figure 5(a) shows two terrain units which join together without stitching. As you can see in the heightfield map of Figure 5(a), as shown in Figure 5(c), the discontinuity is apparent around the joint. Figure 5(b) displays two terrain units which are seamlessly joined together using the minimal cut approach. The corresponding heightfield map of Figure 5(b), as shown in Figure 5(d), reveals no discontinuity around the joint.

4 Experimental Results

We have implemented ant tested the proposed approach on a PC with a processor of Pentium 4 2.4G and 1GB main memory. The graphic card is based on the chipset of GeForce FX 5600. All the codes were written in C# and used DirectX API. There are 698 terrain units segmented from real terrain elevation map of Taiwan in our terrain unit database.

Figure 6 demonstrates heightfield maps of some matching terrain units with respect to the user specified primitives. As you can see in the Figure 6, the matched terrain units are very similar to the given terrain primitives in cross-section, mountain ridge, and terrain contour. Consequently, if a composition of terrain primitives conveys the overall macroscopic terrain in user's mind consistently, the synthesized terrain is intuitively predictable. Also, the resultant synthesized terrain is realistic and reasonable because all terrain units are from real world.

Figure 7(a) demonstrates a synthesized mountain which is composed of two tetrahedron-like terrain primitives, and the corresponding terrain primitives are shown in Figure 7(b) and 7(c). As you can see in Figure 7, the two terrain primitives are used to simulate a cone-like single mountain. Figure 8(a) demonstrates the synthesized mountain range which has one branch and the corresponding terrain primitives are shown in Figure 8(b) and 8(c). It shows that each microscopic terrain feature can be effectively generated by a number of terrain units. Moreover, several microscopic terrain features can be used to synthesize a macroscopic terrain as shown in Figure 9. It shows that the synthesized scene is realistic and flexible.

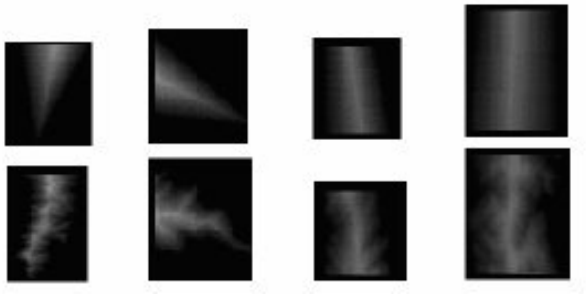
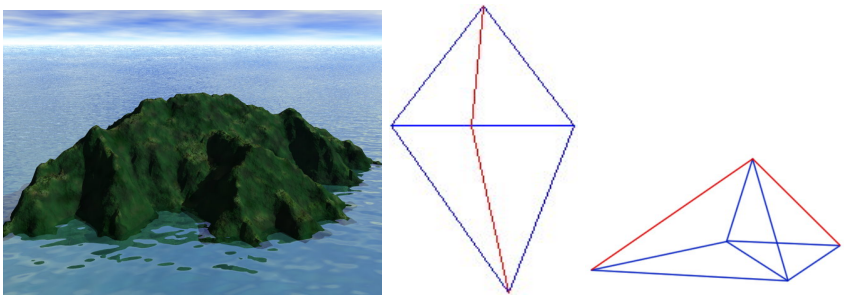
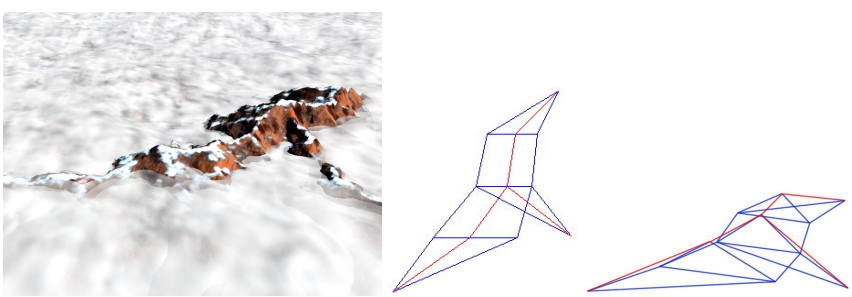


Fig. 6. Terrain primitives and their matched terrain units. Top row: the hightfield map of user specified terrain primitives. Bottom row: the heightfield map of matched terrain units.



(a) (b) (c)

Fig. 7. (a) A mountain is synthesized by two terrain primitives. (b) The top view of the corresponding terrain primitives of (a). (c) A perspective view of the corresponding terrain primitives of (a).



(a) (b) (c)

Fig. 8. (a) A simple mountain range which has one branch is synthesized by several terrain primitives, (b) The top view of the corresponding terrain primitives of (a). (c) A perspective view of the corresponding terrain primitives of (a).

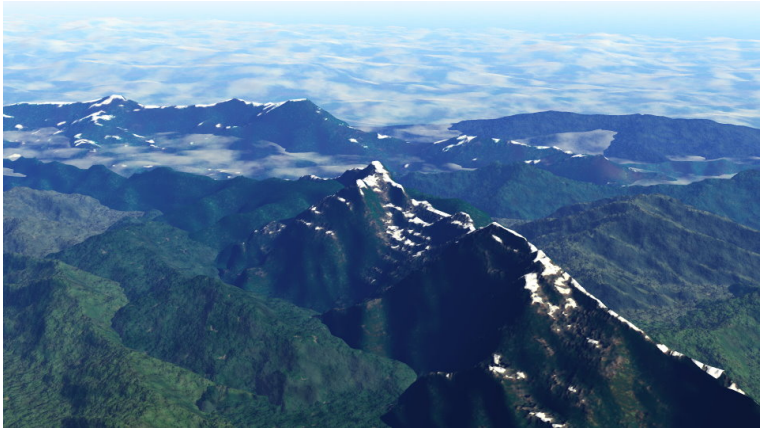


Fig. 9. The synthesized macroscopic terrain using several microscopic terrain features

5 Conclusions and Future Works

In this paper, we use several microscopic feature terrains to synthesize macroscopic terrains. The proposed interactive environment provides an interface that user can intuitively and effectively conceptualize the microscopic terrain features using terrain primitives and then profile the desired macroscopic terrain. The best matching terrain units for the terrain primitives have high geometry and topology similarity. The resultant synthesized terrain is realistic and reasonable because terrain units are from real terrain elevation map and stitched well. Without expensive computation on erosion and simulation used in the previous approaches, our approach is not only effective but also efficient.

Future works include the automatically segmentation method for terrain units and a more flexible stitching method for more than two overlap regions of terrain units. We will exploit the elevation value and the distribution of terrain features to segment terrain units automatically. For seamlessly stitching terrain units, we would try to apply texture synthesis method to synthesize adjacent terrain units with high elevation discrepancy.

References

1. U.S. Geological Survey, http://www.usgs.org/
2. Fournier, A.D., Fussell, D., Carpenter, L.: Computer Rendering of Stochastic Models. Communications of the ACM 25, 338–371 (1982)
3. Gardner, G.Y.: Functional Modeling. In: SIGGRAPH course notes, Atlanta (1998)
4. Lewis, J.P.: Generalized Stochastic Subdivision. ACM Transactions on Graphics 6(3), 167–190 (1987)
5. Mandelbrot, B.B.: Stochastic Models for the Earth's Relief, the Shape and Fractal Dimension of Coastlines, and the Number Area Rule for Islands. Proc. Nat. Acad. Sci. USA 72, 2825–2828 (1975)

6. Mandelbrot, B.B.: The Fractal Geometry of Nature. WH Freeman, New York (1985)
7. Marshall, R., Wilson, R., Carlson, W.: Procedure Models for Generating Three-Dimensional Terrain. Computer Graphics 14(3), 154–162 (1980)
8. Mastin, G.A., Watterberg, P.A., Mareda, J.F.: Fourier Synthesis of Ocean Waves. IEEE Computer Graphics and Applications 7(3), 16–23 (1987)
9. Miller, G.S.P.: The Definition and Rendering of Terrain Maps. Computer Graphics 10(4), 39–48 (1986)
10. Musgrave, F.K., Kolb, C.E., Mace, R.S.: The Synthesis and Rendering of Eroded Fractal Terrains. In: SIGGRAPH 1989, vol. 23(3), pp. 41–50 (1989)
11. Nagashima, K.: Computer Generation of Eroded Valley and Mountain Terrains. The Visual Computer 13, 456–464 (1997)
12. Roudier, P., Peroche, B., Perrin, M.: Landscapes synthesis achieved through erosion and deposition process simulation. Computer Graph Forum 12, 375–383 (1993)
13. Peitgen, H.O., Saupe, D.: The Science of Fractal Images. Springer, New York (1988)
14. Quinlan, S.: Efficient distance computation between non-convex objects. In: IEEE Intern. Conf. on Robotics and Automation, pp. 3324–3329 (1994)
15. Voss, R.F.: Random Fractal Forgeries, Fundamental Algorithms for Computer Graphics. Springer, Heidelberg (1985)
16. Thomas, H.C., Charles, E.L., Ronald, L.R., Clifford, S.: Introduction to Algorithms, 2nd edn. The MIT Press, Cambridge (1990)
17. Vivek, K., Arno, S., Irfan, E., Greg, T., Aaron, B.: Graphcut Textures: Image and Video Synthesis Using Graph Cuts. ACM Transactions on Graphics 22(3), 277–286 (2003)
18. Ford, L., Fulkerson, D.: Flows in Networks. Princeton University Press, Princeton (1962)
19. Chaing, M.Y., Tu, S.C., Huang, J.Y., Tai, W.K., Liu, C.D., Chang, C.C.: Terrain Synthesis: An Interactive Approach. In: International Workshop on Advanced Image Technology, pp. 533–538 (2005)

A Double Domain Based Robust Digital Image Watermarking Scheme

Chuang Lin[1], Jeng-Shyang Pan[2], and Zhe-Ming Lu[3]

[1] Department of Automatic Test and Control, Harbin Institute of Technology, Harbin, P.R. China
linchuang_78@sina.com
[2] Department of Electronic Engineering, National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan
jspan@cc.kuas.edu.tw
[3] School of Information Science and Technology, Sun Yat-Sen University, Guangzhou, P.R. China
luzhem@mail.sysu.edu.cn

Abstract. In this paper, a double domain based robust digital image watermarking scheme is proposed. Up to now, the existing robust watermarking algorithms can only resist a part of geometrical attacks and some common image processing attacks, while the proposed method can resist almost all geometrical attacks and a lot of common image processing attacks. Firstly, the watermark is embedded in the 3-level DWT coefficients by the parity modulation method, and then the same watermark is embedded in the pixel domain by the similar method. At the detector side, the watermark is extracted twice, one is in the DWT domain, the other is in the pixel domain. Select the watermark which has better visual quality as the final result. By carefully regulating the parameters and selecting the embedding sequence, the watermarks in the two domains are compatible with each other. Experimental results show the effectiveness of the proposed method.

Keywords: digital image watermarking, double domain, parity modulation.

1 Introduction

Digital watermarking technique is an important branch of information security. It can be used in copyright protection, authentication, access control, and covert communication etc [1]. Robust watermarking techniques are mainly used in copyright protection etc. In theory, the robust watermark should be resistant to all kinds of attacks. The attacks contain not only common image processing attacks, but also some uncommon attacks, such as geometrical attack. In general, the geometrical attack contains cropping, rotation, enlarging, shrinking, inclining, distortion etc. The common image processing attack includes JPEG compression, Gaussian noise, and filtering etc. In recent years, researchers have proposed many algorithms to resist the geometrical attack. In references [2-5], the authors proposed the Fourier transformation based methods to resist the typical geometrical attacks (rotation,

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 656–663, 2008.

scaling, translation; RST), while the visual quality of the watermarked image is low and they can not resist other geometrical attacks. In references [6], the authors utilizing the techniques which are widely used in image registration to conquer the geometrical attacks, while this kind of methods are very computationally complex and sometimes can not recover the geometrically attacked image to the fine state, and then the watermark cannot be extracted exactly. The performance of the existed watermarking methods which are used to resist the geometrical attacks are very limited by far.

In this paper, a double domains based digital image watermarking scheme is proposed to resist both the geometrical attacks and the common image processing attacks. Firstly, the watermark is embedded in the 3-level DWT coefficients by the parity modulation method, and then the same watermark is embedded in the pixel domain by the similar method. At the detector side, the watermark is extracted twice, one is in DWT domain, the other is in pixel domain. Select the watermark which has better visual quality as the final result. From the experimental results, we can clearly see that, with the same embedding and extracting method, the spatial domain based method has better robustness to the geometrical attacks, but very vulnerable to the common image processing attacks. While the transformation domain based method are on the contrary. By carefully regulating the parameters and selecting the embedding sequence, the watermarks in the two domains are compatible to each other.

The paper is organized as follows: section 2 is the concrete algorithms, section 3 is the experimental results, and section 4 is the conclusions.

2 Embedding the Watermark in the DWT Domain

The detailed embedding procedures in the two domains are a little different. We firstly introduce the watermarking method in the DWT domain.

2.1 Watermark Embedding

Assume the original gray image is A, the binary watermark image is W, $W = (w_{ij})$, $w_{ij} \in \{0,1\}$. (i, j) is a pair of coordinates. The embedding steps are as follows:

Step 1. For A, the 3-level DWT decomposition is performed to get the DWT coefficient set LL_3.

Step 2. For each coefficient p_1 in LL_3, calculate its quantization value λ_1.

$$\lambda_1 = \mathbf{round}(\frac{p_1}{\delta_1}), \tag{1}$$

where **round** is the function which rounds the element to the nearest integer. For example, a = [-1.9, -0.2, 3.4, 5.6, 7.0], **round** [a] = [-2.0, 0, 3, 6, 7]. δ_1 is the quantization step.

Step 3. According to the watermark to be embedded, modify p_1 to p_1^w:

$$p_1^w = \begin{cases} (\lambda_1 - \frac{1}{2})\delta_1, & \lambda_1 + w_{ij} \equiv 1 \pmod 2 \\ (\lambda_1 + \frac{1}{2})\delta_1, & \lambda_1 + w_{ij} \equiv 0 \pmod 2 \end{cases}, \tag{2}$$

Step 4. The IDWT operation is performed to get the watermarked image A_1^w.

2.2 Watermark Extraction

Step 1. For the watermarked image A_1^w, the 3-level DWT decomposition is performed to get the DWT coefficient set LL_3^w.

Step 2. Each DWT coefficient p_1^w in LL_3^w is quantized to λ_1^w.

$$\lambda_1^w = \mathbf{floor}(\frac{p_1^w}{\delta_1}), \tag{3}$$

where **floor** is the function which rounds the element to the nearest integers less than or equal to itself. For example, a = [-1.9, -0.2, 3.4, 5.6, 7.0], **floor** [a] = [-2, -1, 3, 5, 7].

Step 3. Judge the value of the watermark bit w'_{ij} according to the parity of λ_1^w. If λ_1^w is the odd number, the watermark bit w'_{ij} should be 1; if λ_1^w is the even number, the watermark bit w'_{ij} should be 0. That is:

$$w'_{ij} = \begin{cases} 1 & , \lambda_1^w \equiv 1 \pmod 2 \\ 0 & , \lambda_1^w \equiv 0 \pmod 2 \end{cases}, \tag{4}$$

3 Embedding the Watermark in the Pixel Domain

In this part, we introduce the watermarking method in the pixel domain.

3.1 Watermark Embedding

Assume the size of A_1^w is $(mk)\times(nk)$, the binary watermark image is W with size $m\times n$, $W = (w_{ij})$, $w_{ij} \in \{0,1\}$, this watermark is as same as used in DWT domain. The concrete embedding steps are as follows:

Step 1. Divide A_1^w into $m\times n$ blocks with size $k\times k$. The block is denoted as A_{ij}, $A_1^w = (A_{ij})$.

Step 2. For each pixel p_2 in A_{ij}, calculate the quantization value λ_2 of p_2.

$$\lambda_2 = \mathbf{round}(\frac{p_2}{\delta_2}) , \tag{5}$$

where **round** is the function which rounds the element to the nearest integer , and δ_2 is the quantization step.

Step 3. According to the watermark to be embedded, change p_2 to p_2^w .

$$p_2^w = \begin{cases} (\lambda_2 - \frac{1}{2})\delta_2, & \lambda_2 + w_{ij} \equiv 1 \pmod 2 \\ (\lambda_2 + \frac{1}{2})\delta_2, & \lambda_2 + w_{ij} \equiv 0 \pmod 2 \end{cases} , \tag{6}$$

When all the pixels are modified by the above method, we get the watermarked image A_2^w .

3.2 Watermark Extraction

The watermarked image A_2^w may be attacked, so we use A_2^{w*} to represent the attacked watermarked image.

Step 1. Divide A_2^{w*} into pixel blocks with size $k \times k$, the quantity of the pixel blocks is decided by the actual size of A_2^{w*} .

Step 2. Extract the watermark from the pixel blocks. One pixel block corresponds to one watermark bit w'_{ij} . $w'_{ij} = (w')$, w' is the sub-watermark-bit, one pixel in the pixel blocks corresponding to one sub-watermark-bit.

Step 2.1. In each pixel block, calculate the quantization value λ_2^{w*} of pixel p_2^{w*} by the following formula.

$$\lambda_2^{w*} = \mathbf{floor}(\frac{p^w}{\delta_2}) , \tag{7}$$

where **floor** is the function which rounds the element to the nearest integers less than or equal to itself.

Step 2.2. Judge the value of sub-watermark-bit w' according to the parity of λ_2^{w*} .

$$w' = \begin{cases} 1 & , \lambda_2^{w*} \equiv 1 \pmod 2 \\ 0 & , \lambda_2^{w*} \equiv 0 \pmod 2 \end{cases} . \tag{8}$$

If most of the sub-watermark-bit w' is 1, then the watermark bit w'_{ij} is 1; if most of the sub-watermark-bit w' is 0, then the watermark bit w'_{ij} is 0

When all the pixel blocks are processed, the watermark $W' = (w'_{ij})$ is gotten.

4 Experimental Results

In our experiment, the target image is the popular Lena image with size 512×512 and 8 bits-per-pixel resolution. The input watermark is a bird image with size 64×64 and 1 bit-per-pixel resolution, each pixel bit in the watermark image is used as a watermark bit. In the experiment, the value of the quantization step δ_1 and δ_2 is 34.43 and 4.43, respectively. The wavelet base is Haar wavelet. BCR is the abbreviation of bit-correct-rate of the extracted watermark.

The watermark is extracted twice, firstly the watermark is extracted from the pixel domain, and then the watermark is extracted from the DWT domain. Select the watermark which has better visual quality as the final result.

Fig. 1 is the original Lena image, the watermark image, the watermarked Lena image, and the extracted watermark under no attacks. The PSNR value of the watermarked Lena image is 37.4552dB, the watermark is embedded in the DWT and the pixel domain.

(a1) Original Lena image (b1) Watermarked Lena image with PSNR value 37.4552dB

(a2) Original watermark (b2) Extracted watermark, BCR=1

Fig. 1. (a1) Original Lena image. (a2) Original watermark. (b1) Watermarked Lena image. (b2) The extracted watermark.

(a1) the right part is cropped (b1) the underside part is cropped (c1) the surrounding part is cropped

(a2) from the pixel domain (b2) from the pixel domain (c2) from the pixel domain

(a3) from the DWT domain (b3) from the DWT domain (c3) from the DWT domain

Fig. 2. Experimental results under cropping attacks

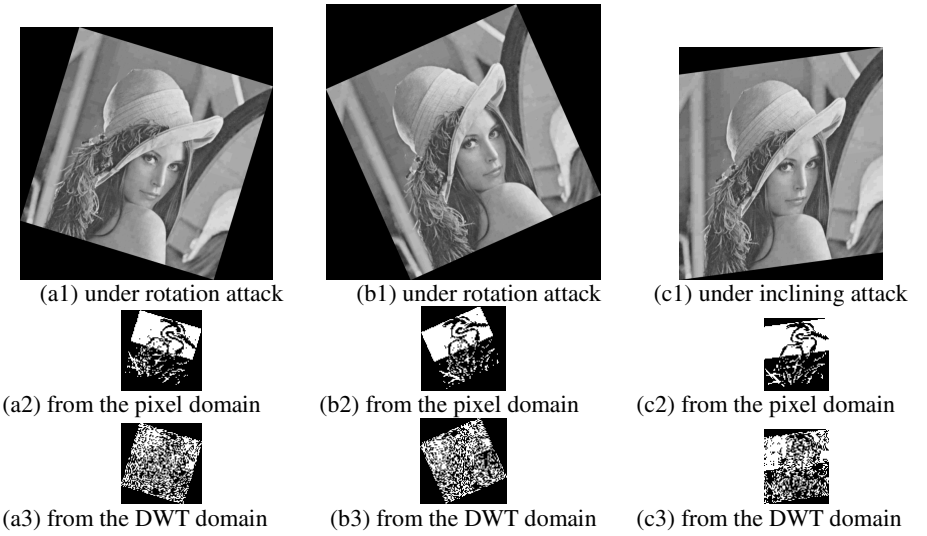


Fig. 3. Experimental results under rotation and inclining, attacks

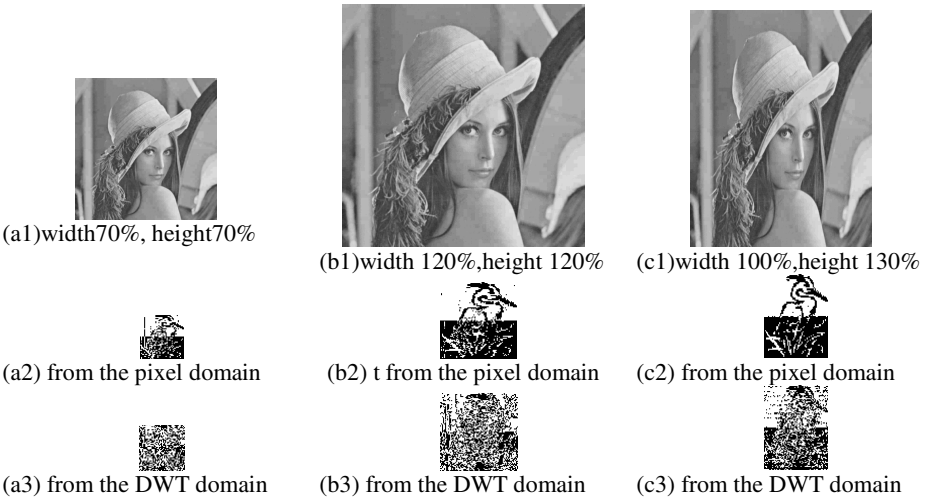


Fig. 4. Experimental results under enlarging and shrinking attacks

Fig. 2 to Figure 4 are the experimental results under geometrical attacks. Fig. 2 is the experimental results under cropping attacks. From the experimental results we can see that, when the watermarked Lena image is cropped, the watermark is also cropped with it. The remained part of the watermark can also be used to insure the copyright. We select the watermark images which have the better visual quality as the final results. Fig. 3 is the experimental results under rotation, inclining, distortion, and perspective attacks. We can see that the extracted watermark is also changed with the

attacked watermarked Lena image. The watermark in DWT domain cannot be extracted under these circumstances. Fig. 4 is the experimental results under enlarging and shrinking attacks. From Fig.1 to Fig. 4, we can get the conclusion that the pixel domain based method is very robust to the geometrical attacks.

Fig. 5 to Fig. 6 are the experimental results under common image processing attacks. Fig. 5 is the experimental results under JPEG attacks with different quality factor (QF). Fig. 6 is the experimental results under noise and filtering attacks. From Fig. 5 and Fig. 6, we can get the conclusion that the DWT based method is very robust to the common image processing attacks.

From the experimental results we can see that, the double domain based watermarking scheme is very robust to the geometrical attacks and the common image processing attacks. Although the method is not so complex, it is effective.

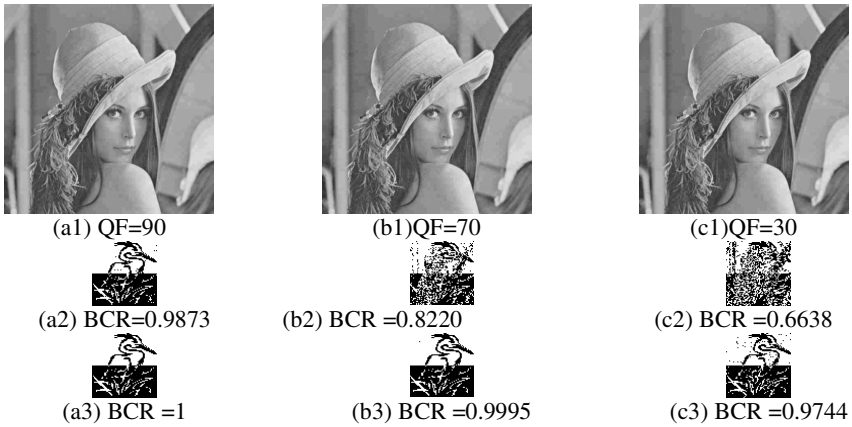


Fig. 5. Experimental results under JPEG compression attack with different quality factor

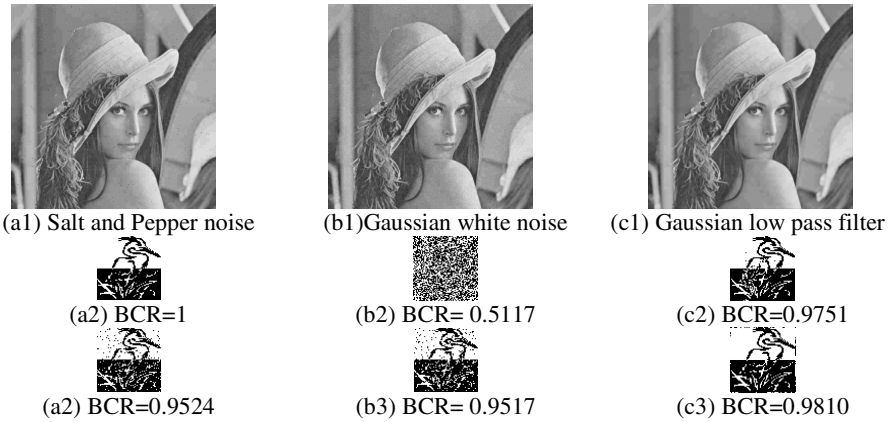


Fig. 6. Experimental results under noise attacks and filter attacks

5 Conclusions

The paper proposed a double domain based watermarking scheme, the concrete embedding and extracting algorithm is fulfilled by the parity modulation method. By carefully regulating the values of the parameters and selecting the embedding sequence, the watermarks in the two domains can smoke the calumet together and make out contributions in different attacks. The experimental results show that the proposed method is very robust to the geometrical attacks and common image processing attacks.

References

1. Yin, H., Lin, C., Qiu, F., et al.: A survey of digital watermarking. Journal of Computer Research and Development 42(7), 1093–1099 (2005)
2. Joseph, J.K., Ruanaidh, Ó., Pun, T.: Rotation, scale and translation invariant spread spectrum digital image watermarking. Signal Processing 66(3), 303–317 (1998)
3. Lin, C.-Y., Wu, M., Bloom, J.A., Cox, I.J., Miller, M.L., Lui, Y.-M.: Rotation, Scale, and Translation Resilient Watermarking for Images. IEEE Trans. on Image Processing 10(5), 767–782 (2001)
4. Pereia, S., Pun, T.: Robust Template Matching for Affine Resistant Image Watermarks. IEEE Trans. on Image Processing 9(6), 1123–1129 (2000)
5. Kang, X., Huang, J., Shi, Y.Q., Lin, Y.: A DWT-DFT Composite Watermarking Scheme Robust to Both Affine Transform and JPEG Compression. IEEE Trans. on Circuits and Systems for Video Technology 8, 776–786 (2003)
6. Braudaway, G.W., Minter, F.: Automatic recovery of invisible image watermark from geometrically distorted images. In: Proceedings on SPIE Security and Watermarking of Multimedia Contents, CA, USA, vol. 3971, pp. 74–81 (2000)

ABF Based Face Texturing

Xia Zhou, Yangsheng Wang, Jituo Li, and Daiguo Zhou

Institute of Automation, Chinese Academy of Science
No.95 Zhongguancun East Road, Room 1208, Beijing, China
{xia.zhou,yangsheng.wang,jituo.li,daiguo.zhou}@ia.ac.cn

Abstract. A new approach of generating textures from multi-view images is presented in this paper. The generated textures can be mapped onto a 3D face model seamlessly. Angle Based Flattening(ABF) surface parameterization is used to build the correspondence between a 3D face model and its 2D texture domain. Feature points on face model are defined according to face anatomy, and their counterparts on images are automatically extracted with AAM method, plus by some little user interaction. The correspondence between feature points on 3D model and 2D images sets up the correspondence between images and the texture domain. With above methods, we can efficiently synthesis texture seamlessly from multi-view images, and map it onto the 3D face model.

Keyword: texture mapping, texture generation, ABF method, AAM.

1 Introduction

Texture mapping is one of the oldest techniques in Computer Graphics [7]. Originally, it is used to convey realism of objects, see [4][5]for an overview . As the methods of texture mapping are throughout computer graphics and image processing [4], its popularity is undoubted. In recent years, it is fast developed and powerfully used in all kinds of modeling not only because it makes the model, like a face, more vivid, but also because the low -cost hardware support for this technique is available. It has already achieved great success in high quality image synthesis [4] and attracts more and more focus from researchers.

In face texturing, there are mainly two kinds of ways to get the texture: 3D scanner, like Cyberware, and image-based texture mapping [5][6].The former one generates both the model and the texture by scanning the real object. It needs a large database which is costly and the generated texture is not so satisfying [3]. A more common way is to use input images. We focus our attentions on these previous works and represent our method in section 1.1..

1.1 Previous Work

In images-based texture generation, we have to solve such a problem: generating a complete, seamless texture from a series of images, and then mapping it on to a 3D face model. Commonly, three or more images are required, and these images are

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 664–674, 2008.

usually unregistered. In[2], Rocchini et al first calibrate a camera by the corresponding feature points in 3D models and images. Then, they create a texture patch for each triangle in 3D model. This is a common way, and deduces to a lot of approaches in the actual implemention of every step. Rocchini's work is effective and can map a texture with lots of details very well, but it can't generate mip-maps as in its texture domain, the textures from all the images are side by side, thus forming a patch structure.

Texture mapping based on parameterization can satisfy the mip-map. It constructs a parameterization for a 3D model over a 2D domain. Then a texture can be created in this parameterization domain. Special parameterizations are used for different kinds of models. In[12], Sheffer et al proposes Matchmaker to improve the parameterization used in face texturing. They embedded a feature mesh (matchmaker) into the original mesh in parameterization result, achieving a low distortion of face model. However, this algorithm is time consuming, and the boundaries are visible in the mapped model. Maroc[1] et al used a view-dependent parameterization to enhance the visual quality of textures, and a multiresolution splines method was applied to remove the boundaries. It works effectively, except that it needs 80 minutes to do the parameterization, and 15-25 minutes more to do the interaction for defining the feature points in 3D model and images.

Most methods of face texturing need lots of manual tuning. To smooth the transitions in the generated texture, additional complex methods are needed after the texture has been generated in texture domain. All above takes lots of time and is not convenient for application. Allowing for these problems, we proposed a new approach to achieve seamless face texturing quickly with few interaction. The paper is organized as follows: First we introduce the anatomy based face feature points detection in section 2. Section 3 is a brief introduction of ABF used in our system. Then, in section 4, we detail the seamless texture generation. The results are shown in section 5. Finally, section 6 gives the conclusion and future work.

2 Anatomy Based Face Feature Point Detection

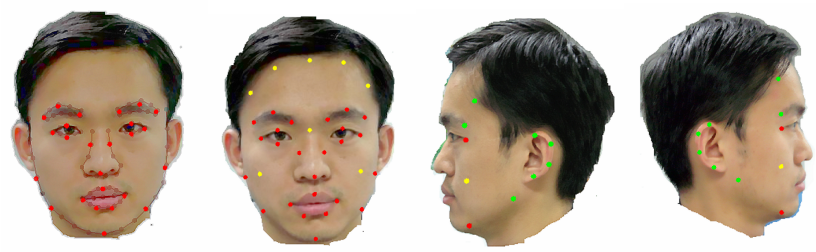
This is a preparing procedure of the texturing. In order to set up the correspondence between 3D face model and its 2D texture domain, we first find the anatomy based face feature points both in 3D model and 2D image domain. Since in surface parameterization, every vertex in 3D face model has a unique counterpart in texture domain, the texture of these feature points can be acquired directly. Then based on this, every 3D vertex can find its correspondence in images, as well as its texture (see details in section 4.1). First of all, the feature points are defined; then, followed by two parts of works: face image feature detection and 3D model feature detection.

2.1 Feature Points Definition

We define 41face feature points. 25 of which are totally anatomy based [16], denoting eyes, mouth, nose and the face contour. In front image, as the triangulation of these 25 feature points can't cover the whole face, we need another 8 more. In each side image, we define 4 more on each ear. See the defined feature points in figure 1 and 2, both in 3D model and 2D image domain.

2.2 Feature Point Detection Both in Face Images and 3Dface Model

To get the complete face texture, we prepare three face images: front image, left and right. In the front face image, we use Active Apparent Model (AAM[14]) to get all 33 feature points automatically. AAM combines face shape and texture information, giving effective and precise face alignment result. In this paper, we use the result of our other work in lab ([15]) directly, selecting 25 of the 87 feature points generated. In both left and right image, only half of the feature points can be seen, we pick 9 in each manually, see figure 1.



(a)AAM result (b)front image (c)left image (d) right image

Fig. 1. (a) Green dots are the AAM results, red dots are selected as our feature points (b) Red dots are from AAM, the yellow ones can be inferred by the 25 features geometrically. (c)(d) Green and red dots are manual, and yellow ones are the middle of reds.

In face model, there are feature points based on anatomy correspondingly. We select them out shown in figure 7:

3 Surface Parameterization

Since surface parameterization can generate mip-mapped texture mapping, it's widely used in this field. A large amount of surface parameterization method has been proposed, for example, Floater[8] uses the solution of a convex combinations based linear system to computer the vertex positions in parameterization plane; Eck et al's[9] proposed harmonic mapping; A. Sheffer presented angle based flattening(ABF) in [10], and so on. Among all the work, we notice that the result of ABF is free boundary and easy to converge. So, we focus on the work on ABF.

3.1 Angle Based Flattening- ABF

ABF is a surface parameterization method. It is an angle-preserving mapping. A Sheffer proposed it based on the observation that for a triangular mesh preserving the size of angles on each of faces is sufficient to maintain the surface metric structures up to a global scaling factor[10].

The core of the algorithm is a minimization problem with three constraints: (see details in [10])

The minimization problem:

$$F(x) = F(\alpha, \lambda_{Tri,} \lambda_{Plan,} \lambda_{Len}) = E + \sum_{t} \lambda_{Tri}^{t} C_{Tri}(t) + \sum_{v} \lambda_{Plan}^{v}(v) C_{Plan}(v)$$
$$+ \sum_{v} \lambda_{Len}^{v}(v) C_{Len}(v), E = \sum_{t \in T} \sum_{k=1}^{3} \frac{1}{w_k^t} (\alpha_k^t - \beta_k^t)^2, w_k^t = \frac{1}{(\beta_k^t)^2} \tag{1}$$

t is the triangle index, k is the angle index in every triangle, α is the required angle in parameterization domain, β is the optimal angle.

The three constraints:

(a)

$$\forall t \in T, C_{Tri}(t) = \alpha_1^t + \alpha_2^t + \alpha_3^t - \pi = 0 \tag{2}$$

(b)

$$\forall v \in V_{\text{int}}, C_{Plan}(v) = \sum_{(t,k) \in v^*} \alpha_k^t - 2\pi = 0 \tag{3}$$

V_{int} is the set of interior vertices, and v^* is the set of angles incident on vertex v.

(c)

$$\forall v \in V_{\text{int}}, C_{Len}(v) = \prod_{(t,k) \in v^*} \sin \alpha_{k \oplus 1}^t - \prod_{(t,k) \in v^*} \sin \alpha_{k \Theta 1}^t = 0 \tag{4}$$

$k \oplus 1$ $k \Theta 1$ indicate the next and the previous angles in the triangles respectively.

See figure 2 to know what the symbols stand for, and the formulas are from [11]

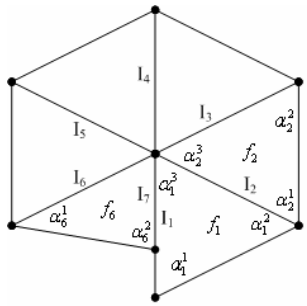


Fig. 2. The brief introduction of ABF

We implement ABF in this paper, and the result can be seen in figure 7. The parameterization result is continuous. In our use, we discretize it into pixels, to be the texture domain. See details in the next section.

4 Seamless Texture Generation from Images

In this section, our goal is to get a seamless texture from multi-view images on parameterization result. In order to do this, we have to know the corresponding pixel pairs in texture domain and the image domain. Then, according to the unique correspondence between 3D face model and its parameterization result, we finish the work of mapping the texture to the face model. We first discretize parameterization result to be pixels, forming the texture domain, then decompose the problem into two as follows:

Step1: find the correspondence of images, face model, and texture domain

Step2: do texture filling and blending in texture domain.

Step 1 is taken to find the matched pixel in images for every vertex in face model. Then, every pixel in texture domain can match a pixel value from images in step 2. After this work, a seamless texture is generated.

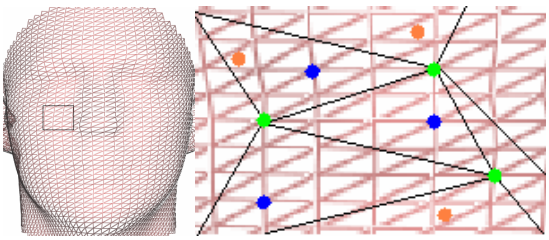
Before we detail each step, we give some definitions, which the whole section is based on.

Definition: In 3D face model, there are feature points vertices and common vertices; In 2D images, there are three kinds of pixels: the feature vertex pixel, the common vertex pixel, and the non vertex pixel; correspondingly in 2D texture domain, there are three kinds of texels: the feature vertex texel, the common vertex texel, and the non vertex texel. We define some symbols to stand for each of them for short, see Table 1:

Table 1. The symbol defined for each kinds of vertexes, pixels and texels

The kind of vertex in 3Dmodel	3D model	2D images	2D texture domain
Feature Points Vertex(Fv)	Fv	p_{Fv}	T_{Fv}
Common Vertex(Cv)	Cv	p_{Cv}	T_{Cv}
Non vertex(Nv)		p_{Nv}	T_{Nv}

Examples are shown in figure 3:



(a)3D face model (b) Texture domain

Fig. 3. (a) The face model. (b) In texture domain, green dots are T_{Fv}, matching Fv in face model and p_{Fv} in images; Blue dots mean T_{Cv}, matching Cv and p_{Cv}; Orange dots are T_{Nv}, matching p_{Nv}.

After parameterization, every Fv is tied to a T_{Fv}, and every Cv find its T_{Cv}. In section 2, p_{Fv} is picked out for every Fv.

We now describe each step in more details.

4.1 Correspondence of Images, Face Model and Texture Domain

In this part, we look for the matched pixel in both images and texture domain for every vertex in face model, which means finding p_{Fv}, T_{Fv} for Fv, and p_{Cv}, T_{Cv} for Cv. T_{Fv}, T_{Cv}, p_{Fv} are already known in the above work, then, so we only have to find p_{Cv} for Cv. We ask T_{Cv} for help and two steps are taken:

Step 1: In texture domain, we find the feature triangle, which T_{Cv} is in, as well as the precise position of T_{Cv} in this feature triangle by barycentric coordinate.

Step 2: Based on the barycentric coordinate, we locate the p_{Cv}.
The barycentric coordinate is used here, see details as follow:

Barycentric Coordinate: For a point and a triangle in the same plane, the coordinates of the point coordinates can be linearly denoted by the three vertices' coordinates, and the coefficients are the barycentric coordinates of the point to this triangle. See figure 4:

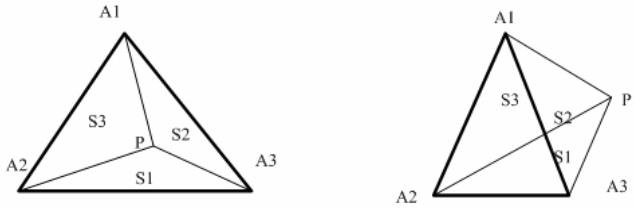


Fig. 4. Barycentric Coordinate

$$p = c_1 A_1 + c_2 A_2 + c_3 A_3, \quad c_i = \frac{s_i}{\sum_{i=1,2,3} s_i} \tag{5}$$

All $c_i > 0$ means P is inside $\triangle A_1 A_2 A_3$; If one of $c_i < 0$, P is outside and we can know it is outside which triangle edge as well, for example, if $c_1 < 0$, P is outside $A_2 A_3$.

Based on this, we can find p_{Cv}. In section 2, feature triangles are formed by feature vertices. For every T_{Cv}, we compute c_i to every feature triangle, find the one with all c_i >0, named $\triangle tri\_f$, and record the c_i as $c_{i\_tri\_f}$ .The $\triangle tri\_f$ has a matched feature triangle in images, we find this triangle and use the same $c_{i\_tri\_f}$ to get the exact p_{Cv}. So, when finding the $\triangle tri\_f$, p_{Cv} is also found.

However, this way to find the $\triangle tri\_f$ is low efficient obviously. We used an improved way instead. See figure 5. We compute the c_i of a T_{Cv} to an arbitrary feature triangle. If all the c_i >0, $\triangle tri\_f$ is found; else, using c_i, we still can know T_{Cv} outside which edge , and tag it edge E. Then we compute the c_i of T_{Cv} to the feature triangle sharing edge E with the first feature triangle. Repeating the operation until finding the $\triangle tri\_f$. The way works effectively in the experiment.

Obviously, to each face model vertex, the corresponding pixel in images is not unique. There may be two pixels, one in front image, the other in one side image, tied to a vertex. We use an area-weighted method to handle with this situation. See details in section 4.2.

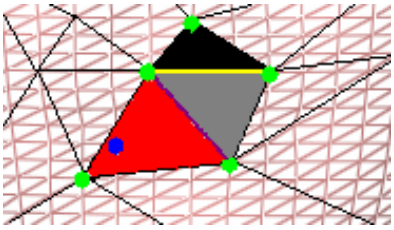


Fig. 5. Find a T_{Cv} 's location in texture domain: black, gray and red triangles are all feature triangles. Compute c_i of T_{Cv} to black triangle, and know T_{Cv} is outside yellow edge, then find the grey one. Repeat the operation, then we find the $\triangle tri\_f$ for the T_{Cv}, that is the red triangle in the figure.

4.2 The Texture Filling and Blending on Parameterization Result

In this section, we first attach pixel values to T_{Fv} and T_{Cv} from multi-view images, then set the pixel values to T_{Nv}.

For T_{Fv} and T_{Cv}, the corresponding p_{Fv} and p_{Cv} are already found in multi-view images after section 4.1. We use an area-weighted method to handle with the situation that one texel in texture domain has two matched pixels in images. For example, an T_{Fv} has one matched pixel in front image, $p_{Fv}1$, in feature triangle $\triangle F$, another matched pixel in left image, $p_{Fv}2$, in $\triangle L$. S stands for the triangle's area. We

set $\alpha = s_{\triangle F} / (s_{\triangle F} + s_{\triangle L})$ $\beta = s_{\triangle L} / (s_{\triangle F} + s_{\triangle L})$. Then $T_{Fv} = \alpha T_{Fv}1 + \beta T_{Fv}2$. This method uses the texture information in multi-view images sufficiently and can explain the details. It also helps smooth transition between images to achieve seamlessness.

All the T_{Cv} form lots of triangles $\triangle para$ in texture domain, each corresponding to a mesh triangle in model. There are a large amount of discrete T_{Nv}, and we even don't know the correspondence of T_{Nv} and these $\triangle para$. In order to get the corresponding p_{Nv} in images, we traversal every $\triangle para$, setting some sample T_{Nv} in it. See figure 6:

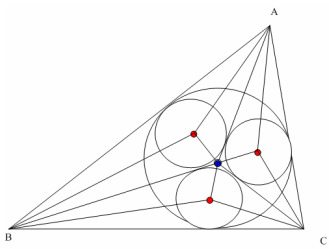


Fig. 6. The blue point is the first selected incenter, the red ones are the second selected

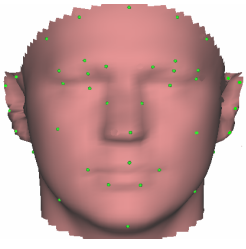
We choose the incenter of the $\triangle para$ as the first selected sample. This incenter separates the $\triangle para$ into three triangles. The second selected samples are the incenters of the three small triangles. Four times recursion generate 40 T_{Nv}, which is proved enough to represent the $\triangle para$ in our experiment. Then the pixel value of left T_{Nv} in the $\triangle para$ can be set the same as its adjacent sample T_{Nv}.It works fast, and the result turns to be legible. See result in Figure 7.

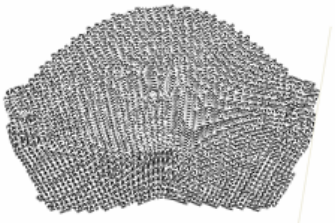
We discretize the parameterization result into pixels, and copy the pixel values from multi-view images one by one. This operation minimizes the norm effect on texture, and generates a seamless texture, without considering the illumination variation.

5 Results

We experimented on several individuals. All the experiments are done on a common computer with PC 1.7 GHz Pentium 4.The 3D face model feature points detection is a pure interactive step, taking about 5 minutes. But we need to do this only once. It is a step in system design, not in the application. In 2D image feature detection, as AAM is real-time, left and right side images interaction cost all of the time in this step, about 2 minutes each. ABF takes 4 seconds for a mesh of 4096 triangles. Getting the pixel value for every pixel in the texture domain (616×496 pixels) from images needs 12 seconds. The whole procedure costs about 4.25 minutes from inputting three images to getting the textured face. See the result in figure 7:

(a) Three input images





(b) The face model with feature points(c) The parameterization result

(d) The generated texture on parameterization result (the texture domain)

(e) The textured face

Fig. 7. (a)The three input images (b)The face model(c)the ABF result(d)The generated texture on parameterization result(e)The textured face

6 Conclusion and Future Work

We introduced a new approach to seamlessly texture face. A popular face alignment algorithm AAM is used to detect 2D face feature points, which helps reduce time and interaction. According to the result of [15], we need some little interactions in side images. To do the seamless texturing, we implement ABF to flatten the 3D face model first, then discretize the parameterization result to be the texture domain, and copy the pixel value from the images one by one. The texture generated in this way is seamless unless the face texture of three input images differs obviously. This work will be used to texture the face model in[13], and also be used in face animation.

We could make several improvements that can be made in future . As AAM was originally used to detect the feature point in computer vision, all the points detected are on the face, with no ear feature points. For texture mapping, we can extend AAM to detect point on ears as well, based on which, we can infer the ear's feature points in side images. This helps to reduce the interaction, and the time can be saved greatly. In the texture domain, we can see the distortions when we "move" the pixel value from images. We consider some constrains from the images, and add it to improve ABF used in texturing face.

References

1. Tarini, M., Yamauchi, H., Haber, J., Seidel, H.-P.: Texturing Faces. In: Proc. of Graphics Interface, Calgary, Canada, pp. 89–98 (2002)
2. Rocchini, C., Cignoni, P., Montani, C., Scopigno, R.: nultiple Textures Stitching and Blending on 3D Objects. In: Proceedings of Eurographlcs Workshop on Rendering, Granada, pp. 173–180 (1999)
3. Blanz, V., Vetter, T.: A Morphable Model For The Synthesis Of 3D Faces. In: Proceedings of ACM SIGGRAPH99 Conference, Los Angeles, pp. 187–194 (1999)
4. Heckbert, P.S.: Survey of Texture Mapping. IEEE Comput. Graph. Appl. (Sept.), 56–67 (November 1986)
5. Ebert, D.S., Musgrave, F.K., Peachey, D., Perlin, K., Worley, S.: Texturing and Modeling: A Procedural Approach, 2nd edn. Academic Press, London (1998)
6. Yamauchi, H., Lensch, H.P.A., Haber, J., Seidel, H.-P.: Textures revisited. The Visual Computer 21(4), 217–241 (2005)
7. Blinn, J.F., Newel, M.E.: Texture and Reflection in Computer Generated Images. Communications of the ACM 19(10), 542–546 (1976)
8. Floater, M.S.: Parameterization and Smooth Approximation of Surface Triangulations. Computer Aided Geometric Design 14, 231–250 (1997)
9. Eck, M., DeRose, T., Duchamp, T., et al.: Multiresolution Analysis of Arbitrary Meshes. In: Computer Graphics. SIGGRAPH's 1995. Annual Conference Series, pp. 173–182 (1995)
10. Sheffer, A., de Sturler, E.: Parameterization of Faceted Surfaces For Meshing Using Angle-Based Flattening. Engineering with Computers 17(3), 326–337 (2001)
11. Sheffer, A., Lévy, B., Mogilnitsky, M., Bogomyakov, A.: ABF++: Fast and Robust Angle Based Flattening. ACM Transactions on Graphics (TOG) 24(2), 311–330 (2004)
12. Kraevoy, V., Sheffer, A., Gotsman, C.: Matchmaker: Constructing Constrained Texture Maps. ACM Transactions on Graphics 22(3), 326–333 (2003)

13. Ding, B., Wang, Y., Yao, J., Lu, P.: Fast Individual Face Modeling and Facial Animation System. In: The 1st International Conference of E-Learning and Games, Hangzhou, China (April 2006)
14. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active Appearance Models. In: Proc. European Conf. Computer Vision 2, pp. 484–498 (1998)
15. Wang, S., Wang, Y., Chen, X.: Weighted Active Appearance Models. In: Huang, D.-S., Heutte, L., Loog, M. (eds.) ICIC 2007. LNCS, vol. 4681, pp. 1295–1304. Springer, Heidelberg (2007)
16. EDU Resources.NET: http://www.eduresources.net/science/anatomy/bsh00.htm

Tile-Based Interactive Texture Design

Weiming Dong[1], Ning Zhou[2], and Jean-Claude Paul[2,3]

[1] LIAMA-NLPR, CAS Institute of Automation, Beijing, China
wmlake@gmail.com
[2] Tsinghua University, Beijing, China
[3] INRIA, France

Abstract. In this paper, we present a novel interactive texture design scheme based on the tile optimization and image composition. Given a small example texture, the design process starts with applying an optimized sample patches selection operation to the example texture to obtain a set of sample patches. Then a set of ω-tiles are constructed from these patches. Local changes to those tiles are further made by composing their local regions with the texture elements or objects interactively selected from other textures or normal images. Such select-compose process is iterated many times until the desired ω-tiles are obtained. Finally the tiles are tiled together to form a large texture. Our experimental results demonstrate that the proposed technique can be used for designing a large variety of versatile textures from a single small example texture, increasing or decreasing the decreasing the density of texture elements, as well as for synthesizing textures from multiple sources. Keywords: Interactive texture design, Texture synthesis, Image composition, ω-tile

1 Introduction

Textures have been a research focus for many years in human perception, computer graphics and computer vision. Recently, research activities in this area emphasize on texture synthesis. Given an example texture, a texture synthesis algorithm generates a new one bearing the same visual characteristics. In spite of the fact that numerous methods have been proposed for texture synthesis, how to design a variety of large textures from a single small example texture is still a challenging problem.

Recently, Matusi et al. [1] developed a system for designing novel textures in the space of textures induced by an input database. However, their texture interpolation technique is based on a single one-to-one warping between the pairs of texture examples, which might be too restrictive for textures with highly irregular structures, causing discontinuous mappings of the patches to the original image. Shen et al. [2, 3] proposed deformation-based texture design techniques for producing a variety of textures by applying deformations and energy optimizations to the extracted patches of texture elements. The main limitations of Shen et al.'s methods, lie in the facts that they have no reusability for the resulting texture elements and the run-time synthesis speed is low.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 675–686, 2008.

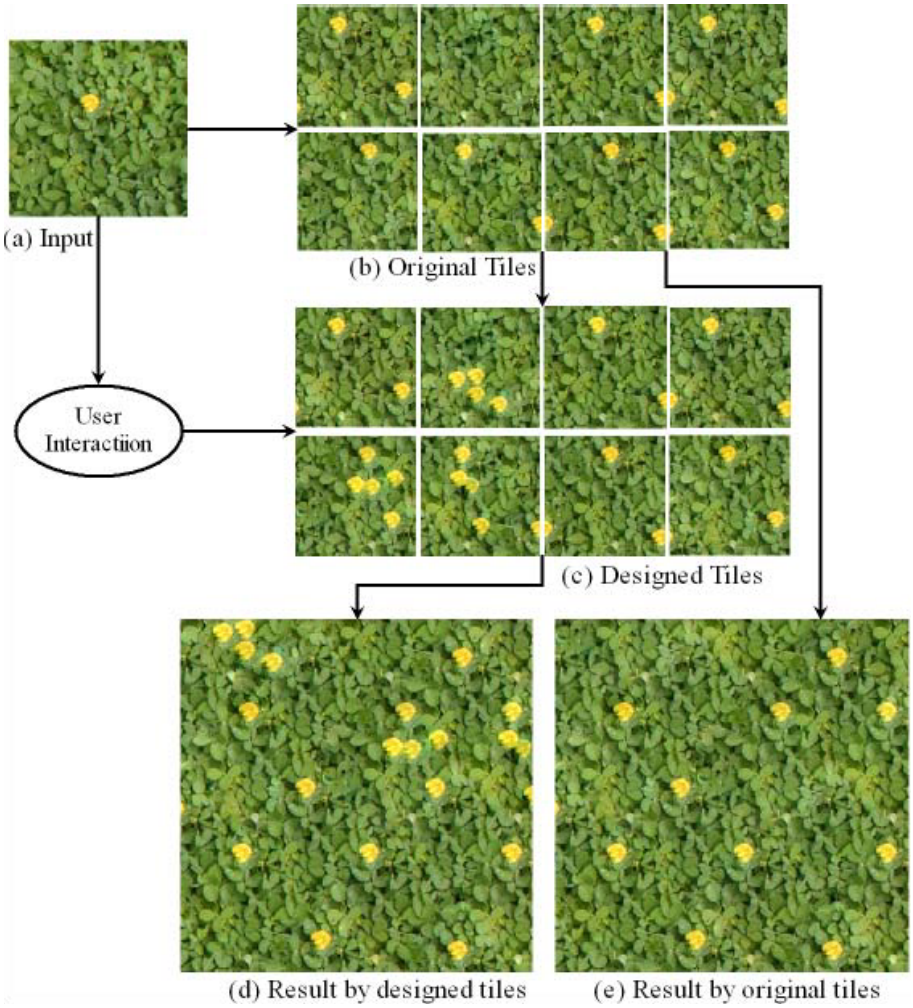


Fig. 1. Our tile-based interactive texture design algorithm. The texture element of the yellow flower is added into some tiles by user interactions. We can see that the designed tiles can increase the variations of the result textures.

The tile-based texture synthesis technique is to use texture synthesis to precompute a set of small texture tiles and use these tiles to generate arbitrary size of non-periodic images at run time [4–8]. The tile-based method usually employs a set of sample patches which are extracted from the input example as texturing primitive. Then tiles are constructed by stitching sample patches together following some given rules. The technique requires only a small amount of memory and is very useful in many real-time applications. We use ω-tiles [6, 8] as the tile patterns in our system.

In this paper, we present a new tile-based interactive texture design algorithm. The proposed algorithm has the ability to locally change the visual property of texture tiles

with little user interaction, and hence drastically broadens the variations of textures that can be synthesized with the existing tile-based methods. As shown in Fig. 1, from a single small example texture, our technique can create a variety of versatile textures, with increased or decreased density of texture elements in the tiles. The main contributions of our work consist of the following three aspects:

(1) A novel framework for designing a large variety of textures by integrating the techniques of 1) tile-based texture synthesis, 2) interactive image editing, 3) genetic algorithm (GA) based optimization, and 4) gradient-based Poisson image composition.
(2) An effective GA-based method for automatically extracting optimized sample patches from the input example texture.
(3) A new composition based algorithm for synthesizing texture tiles from multiple sources.

In the rest of the paper, we first introduce the related work on texture synthesis and interactive image manipulation tools in Sec. 2. Then, in Sec. 3, we discuss the details of our tile-based interactive texture design scheme. The extension of the existing tile optimization algorithm using genetic algorithm is also described in Sec. 3. The method for synthesizing textures from multiple sources using Poisson image composition is presented in Sec. 4. After showing the experimental results in Sec. 5, we conclude the paper and show some directions for future work in Sec. 6.

2 Related Work

2.1 Texture Synthesis

Nowadays local region-growing methods are popularly used in texture synthesis. These methods generate the texture by growing one pixel [9–12] or patch [13–18] at a time with the constraint of maintaining coherence of neighboring pixels in the grown region. Such approaches always suffer the time-consuming neighborhood matching in the example and do not sufficiently meet real-time applications. On the other hand, some near real-time texture synthesis methods usually achieved low quality results for the lack of optimization in the pre-processing [19, 20] or needed very complex pre-computation and data structures [21]. Recently efficient GPU-based texture synthesis techniques [22, 23] have also been proposed, however they always demand a high performance graphics hardware, and their methods suffer from the pixel-based synthesis issue of performing poorly on textures with semantic structures not captured by small neighbor-hoods.

An alternative approach is the tile-based methods. Cohen et. al. [4] developed a stochastic algorithm to non-periodically tile the plane with a small set of Wangtiles at run time. Wei [5] extended this work with GPU to improve tile-based texture mapping. Ng et al. [6] presented another approach to generate a set of small texture tiles from an input example. These tiles could also be tiled together to synthesize large textures. Dong et. al [8] extend the algorithm in [6] to derive new tile sets by increasing the number of sample patches. Our technique uses their ω-tiles as the tile set pattern.

2.2 Interactive Image Manipulation Tools

Interactive image manipulation and editing packages, such as Adobe Photoshop, are commonly utilized by digital photographers. In their workflow, images are manipulated directly and immediate visual feedback is provided.

Recently, many researchers proposed several interactive digital image editing tools by using region-based methods, e.g., the magic wand in Photoshop, intelligent paint [24], interactive graph-cut image segmentation [25], Poisson image editing [26], GrabCut [27], lazy snapping [28], interactive image photomontage [29], drag-and-drop pasting [30] and photo clip art [31].

The tile design process in our framework is most closely related to the method of drag-and-drop pasting [30], where users use brushes to indicate which parts of a set of photographs should be combined into a composite tile. By allowing the user to interact with local texture elements, our proposed algorithm has the ability to increase or decrease the density of texture elements interactively, which is suitable for designing a variety of versatile textures from a single small example texture.

3 Our Approach

3.1 Algorithm Overview

The goal of our algorithm is to enable the texture designer to easily create a set of ω-tiles, from single or multiple sources.

Our proposed work flow is summarized as follows:

1. Load a small example texture image I .
2. Select a group of sample patches from I . The number of sample patches is specified by the pattern of the ω-tile set which is going to be used [8]. Then construct the tile set $T_{o1}, T_{o2}, \cdots, T_{om}$ with the extracted sample patches, as described in Sec. 3.2.
3. Design the ω-tiles by copying texture elements or other objects from the input sources.
4. Repeat Step (3) until a satisfactory set of tiles $T_{d1}, T_{d2}, \cdots T_{dm}$ is obtained.
5. Synthesize the final results by texture tiling [6, 8].

This work flow is illustrated by the sequence of images in Fig. 1. Given an input example texture (Fig. 1(a)), a set of ω-tiles are produced (Fig. 1(b)) after applying the sample patches selection and tile construction operations. Then the user uses brushes to paint some texture elements interactively and the corresponding regions of those texture elements are automatically calculated. These texture elements are stitched into the tiles with the gradient-based Poisson image composition [26, 30], in order to obtain the textures with varying local properties (Fig. 1(c)). Finally, by tiled the ω-tiles together, large textures are synthesized (Fig. 1(d)).

3.2 Optimized ω-Tile Construction

In order to make the above work flow effective, several requirements should be met, such as quickly generated previews of the overall result, a simple, intuitive and easy to

use mechanism for performing the local modification, and an undo function allowing the user to modify previously specified adjustments. Our prototype implementation is based on the interactive digital photomontage technique [29] and the deformation-based interactive texture design system [3]. Besides the above requirements, another important aspect we should notice is the construction of the original ω-tile set. We should assure to generate a high quality tile set. For this, the most important step is the selection of sample patches [8].

We use a similar framework of Dong et al. [7, 8] to select a set of optimized sample patches from the input example. The algorithm is essentially based on genetic algorithm (GA) [32–36]. GA starts with an initial set of randomly generated chromosomes called a population where each chromosome encodes a solution to the optimization problem. All chromosomes are evaluated by an evaluation function which is some measure of fitness. A selection process based on the fitness values will form a new population. The cycle from one population to the next is called a generation. In each new generation, all chromosomes will be updated by the crossover and mutation operations. Then the selection process selects chromosomes to form a new population. After performing a given number of cycles, or other termination criteria is satisfied, we denote the best chromosome into a solution, which is regarded as the optimal solution of the optimization problem.

The main limitation of GA is sometimes it plunges the objective into a local optimal solution. So we improve the GA-based optimized sample patches selection framework in [8] by adding a "genetic-dominance obtaining strategy" before the crossover process of GA. The concept of our technique is as similar as the immuno-dominance obtaining strategy in artificial immune system [37–39]. We denote the sets of sample patches (chromosomes) as $\mathrm{A} = [A_1, A_2, \cdots A_k]$,

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kn} \end{pmatrix}$$

We define a reference chromosome $c = [c_1, c_1, \cdots c_n]$, where

$$k_i = \begin{cases} 1, & \frac{1}{n}\sum_{j=1}^{n} a_{,jj \geq 0.25} \\ 0 & otherwise \end{cases}$$

Denote $A_i \in A$ is the best chromosome in the population, if $F(c) > F(A_i)$, then exchange each other. Note here F is the fitness function of the genetic algorithm [8].

We make $A_j \in A,\ j = 1, 2, ..., k,\ j \neq i$ to get the genetic-dominance referring to a possibility p_{id}. Specifically, we set $a_j^{'} = H(a_j + a_i - c - 1)$, where

$$H(x) = \begin{cases} 1 & x > 0 \\ 0 & otherwise \end{cases}$$

We could adaptively adjust the value of p_{id} . For example, if $F(a_j^{'}) > F(a_j)$ or $F(c) > F(a_i)$), it means that the genetic-dominance is affective, then we increase p_{id} , otherwise we can decrease it.

The genetic-dominance technique combine the global search and local search together, and also import the prior knowledge and the strategy of adaptively obtaining prior knowledge. On the other hand, it achieves the information exchange between the individuals (different with the crossover operation in GA), so it assures the diversity of the population and increases the efficiency of the whole algorithm.

We also add the feature mask [16] to the neighbor matching step of the sample patches selection operation. Given a user-provided binary feature mask, we include it as an additional image channel prior to the neighborhood analysis in [8].

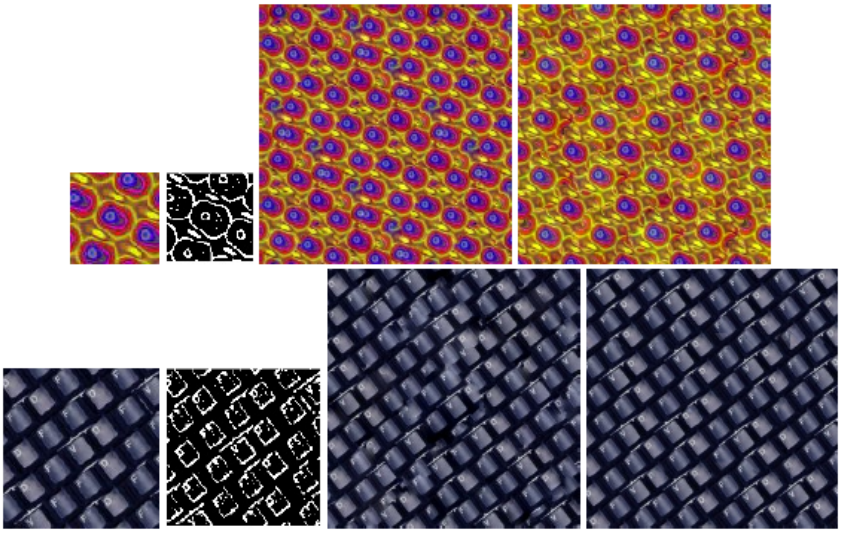


Fig. 2. Compariation of the texture synthesis results using our algorithm and the algorithm in [8]. Form left to right: input example, feature mask, the result using the algorithm in [8], our results.

As shown in Fig. 2, we can see that our algorithm can generate better results than [8] for the examples with some structural texture elements.

3.3 Interactive Local Texture Design

At Step (3) of our algorithm, the local change of an ω-tile is realized by replacing its local regions with the texture elements from the input example. We call the ω-tile to be locally designed the *base tile* $T_{base}(T_{base} \in \{T_{b1}, T_{b2}, ..., T_{bm}\})$ and the source providing the texture elements the *reference texture* I_{ref} ($I_{ref} = I$ in our experiments). As shown in Fig. 1, the user does not need to precisely specify the region including the texture

elements (”yellow flowers”) in I_{ref} . The corresponding region including the texture elements is calculated automatically with the graph-cut-based energy optimization technique [14]. The obtained texture elements are then embedded into to the base tile *Tbas*e seamlessly by the gradient-based Poisson optimization method [26, 30]. Such local design is repeated several times, while at each step the user is allowed to choose new texture elements by painting new strokes according to his creation.

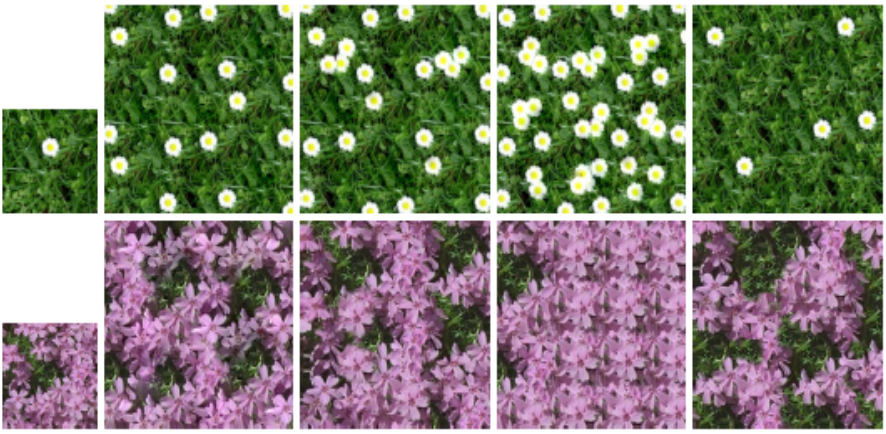


Fig. 3. Examples of spatially varying designed textures using our tile-based method. Left two columns are the input textures and results without tiles design, the others are the synthesized textures with tiles design.

4 Texture Design from Multiple Sources

Our texture design method from multiple sources using image composition is extend from [30]. The goal of multi-source texture design is to synthesize new textures that capture the combined characteristics of several input images. As illustrated in Fig. 4, a desert texture is selected as the background. This kind of textures can be frequently explored in many computer games. Without losing generality, we use the basic 8-tiles ω-tile set in [6, 8] as the bearer in our demos so that more ”designed” patterns can be shown in a normal size tiling. To embed objects in some tiles, we first use GrabCut [27] to produce the rough boundaries of the interested objects from the source images, then the algorithm in [30] is proceeded to find the optimized boundaries. Finally we simply use Poisson image composition to embed the objects in the specific tiles, as shown in Fig. 4(b). In Fig. 4(c), we can see an image which illustrates a desert with pyramids, stones, desert plants and sand dunes. It is generated in real-time with the tiles in Fig. 4(b) and will be more vivid than a simple bare desert when it appears in a computer game. We can choose ω-tile set with more tiles in it [8] to construct designed tiles when

synthesizing large size images, especially for certain applications where continuous patterns are required.

For some objects of the input sources, the sizes may be too large to be embedded in a single tile of smaller size. In this case, we can compose the interested object with multiple tiles, i.e. with a "small" tiling of the tile set. Fig. 6 shows a demo of this instance, here we embed Tintin with 4 tiles which can be tiled together to form a relatively larger background. Then in the tiling process, we just use the Tintin pattern when the same tiles tiling sequence is detected.

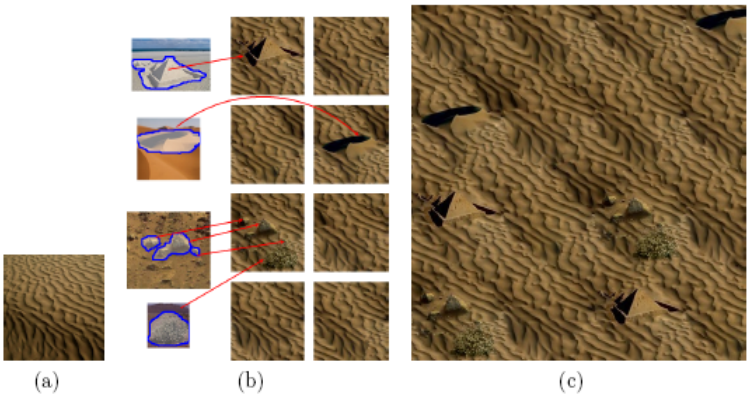


Fig. 4. A desert with pyramids, stones, desert plants and sand dunes. (a) Background example texture. (b) Tiles design with poisson image editing from multiple sources. The blue curves are the rough boundaries of the objects. (c) Image synthesis.

5 Results and Discussions

Our algorithm has been applied to a variety of example texture images. All the experiments shown in this paper were run on a PC with Core2 Duo E6550 2.33GHz CPU + 2GB RAM.

In Fig. 2, we compare our approach with the existing tile-based texture synthesis technique in [8]. The synthesis quality of some structural texture examples could be effectively improved using our enhanced GA-based ω-tile construction algorithm.

Fig. 3 gives examples demonstrating the capability of our technique for creating a large variety of textures from a small examples, while maintaining the continuity of texture features as well as the shapes of individual texture elements. Our methods changes the density of texture elements (white flowers in the first row, the flowers and grass in the second row) interactively according to the designer's need. Fig. 5 shows another result of image synthesis with tiles design using multiple sources. It is a busy water surface with bears, water birds and boat models. We can see that our technique can generate versatile images with textural backgrounds.

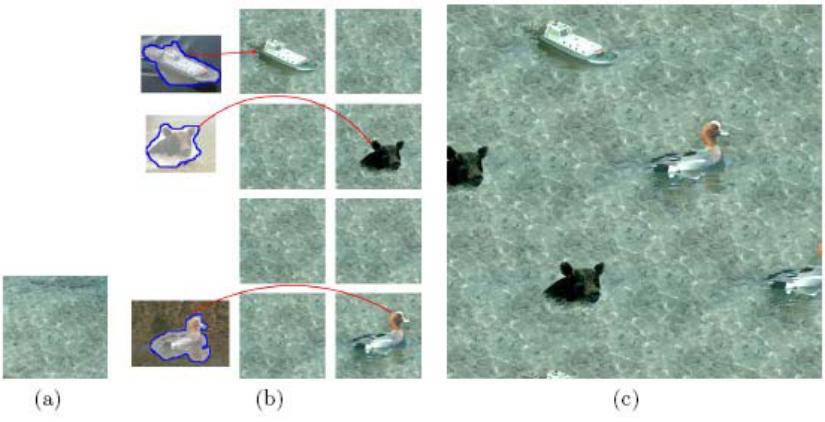


Fig. 5. Water surface with bears, water birds and boat models. (a) Background example texture. (b) Tiles design. (c) Real-time tiling.

Fig. 6. A stone wall painted with Tintin, Doreamon and Mickey Mouse. The image is generated in real-time with our tile-based interactive texture design algorithm.

6 Conclusions and Future Work

A novel tile-based interactive texture design method using image composition has been proposed in this paper. Experimental results demonstrate both the feasibility and the effectiveness of our algorithm. Our algorithm can create a wide variety of very natural textures interactively, from a single small example texture, according to the texture designer's need and creation. The main advantage of our technique over most existing texture synthesis and texture design methods lies in its capability to construct reusable tiles as a pre-computation and generate arbitrary size images in real-time at run time. It is very difficult for the local region-growing-based texture synthesis or existing texture design methods as [2, 3]. Our experimental results also demonstrate that the proposed technique can be applied to other applications such as image synthesis from multiple sources. This approach is very useful in many fields such as interactive decorative pictures design and land map generation of computer games.

Although the tile design operations used in our method can produce good results, it would be meaningful to develop more sophisticated and powerful texture design tools in the future. On the other hand, we would also like to improve the interactive form to make the tool more convenient and automatic.

Acknowledgement

We thank Jiaya Jia, Vivek Kwatra and Sylvain Lefebvre for sharing their results and texture examples on the web sites. This work is supported by National Natural Science Foundation of China projects No. 60073007, 60473110; by National High-Tech Research and Development Plan 863 of China under Grant No. 2006AA01Z301; and by the MOST International collaboration project No. 2007DFC10740.

References

[1] Matusik, W., Zwicker, M., Durand, F.: Texture design using a simplicial complex of morphable textures. ACM Trans. Graph. 24(3), 787–794 (2005)

[2] Shen, J., Jin, X., Mao, X., Feng, J.: Completion-based texture design using deformation. The Visual Computer 22(9-11), 936–945 (2006)

[3] Shen, J., Jin, X., Mao, X., Feng, J.: Deformation-based interactive texture design using energy optimization. The Visual Computer 23(9-11), 631–639 (2007)

[4] Cohen, M.F., Shade, J., Hiller, S., Deussen, O.: Wang tiles for image and texture generation. ACM Trans. Graph. 22(3), 287–294 (2003)

[5] Wei, L.Y.: Tile-based texture mapping on graphics hardware. In: HWWS 2004: Proceedings of the ACM SIGGRAPH/EUROGRAPHICS conference on Graphics hardware, pp. 55–63. ACM Press, New York (2004)

[6] Ng, T.Y., Wen, C., Tan, T.S., Zhang, X., Kim, Y.J.: Generating an ω-tile set for texture synthesis. In: Proceedings of Computer Graphics International 2005 (CGI 2005), Stone Brook, NY, USA, pp. 177–184 (2005)

[7] Dong, W., Sun, S., Paul, J.C.: Optimal sample patches selection for tile-based texture synthesis. In: CAD-CG 2005: Proceedings of the Ninth International Conference on Computer Aided Design and Computer Graphics (CAD-CG 2005), pp. 503–508. IEEE Computer Society, Washington (2005)

[8] Dong, W., Zhou, N., Paul, J.C.: Optimized tile-based texture synthesis. In: GI 2007: Proceedings of Graphics Interface 2007, pp. 249–256. ACM, New York (2007)
[9] Bonet, J.S.D.: Multiresolution sampling procedure for analysis and synthesis of texture images. In: SIGGRAPH 1997: Proceedings of the 24th annual conference on Computer graphics and interactive techniques, pp. 361–368. ACM Press/Addison-Wesley Publishing Co., New York (1997)
[10] Efros, A.A., Leung, T.K.: Texture synthesis by non-parametric sampling. In: ICCV 1999: Proceedings of the International Conference on Computer Vision, vol. 2, p. 1033. IEEE Computer Society, Washington (1999)
[11] Wei, L.Y., Levoy, M.: Fast texture synthesis using tree-structured vector quantization. In: SIGGRAPH 2000: Proceedings of the 27th annual conference on Computer graphics and interactive techniques, pp. 479–488. ACM Press/Addison-Wesley Publishing Co., New York (2000)
[12] Ashikhmin, M.: Synthesizing natural textures. In: SI3D 2001: Proceedings of the 2001 symposium on Interactive 3D graphics, pp. 217–226. ACM Press, New York (2001)
[13] Efros, A.A., Freeman, W.T.: Image quilting for texture synthesis and transfer. In: SIGGRAPH 2001: Proceedings of the 28th annual conference on Computer graphics and interactive techniques, pp. 341–346. ACM Press, New York (2001)
[14] Kwatra, V., Schodl, A., Essa, I., Turk, G., Bobick, A.: Graphcut textures: image and video synthesis using graph cuts. ACM Trans. Graph. 22(3), 277–286 (2003)
[15] Nealen, A., Alexa, M.: Hybrid texture synthesis. In: EGRW 2003: Proceedings of the 14th Eurographics workshop on Rendering, Aire-la-Ville, Switzerland, Switzerland, Eurographics Association, pp. 97–105 (2003)
[16] Wu, Q., Yu, Y.: Feature matching and deformation for texture synthesis. ACM Trans. Graph. 23(3), 364–367 (2004)
[17] Liu, Y., Lin, W.C., Hays, J.: Near-regular texture analysis and manipulation. ACM Trans. Graph. 23(3), 368–376 (2004)
[18] Kwatra, V., Essa, I., Bobick, A., Kwatra, N.: Texture optimization for example-based synthesis. ACM Trans. Graph. 24(3), 795–802 (2005)
[19] Zelinka, S., Garland, M.: Towards real-time texture synthesis with the jump map. In: EGRW 2002: Proceedings of the 13th Eurographics workshop on Rendering, Aire-la-Ville, Switzerland, Eurographics Association, pp. 99–104 (2002)
[20] Zelinka, S., Garland, M.: Jump map-based interactive texture synthesis. ACM Trans. Graph. 23(4), 930–962 (2004)
[21] Liang, L., Liu, C., Xu, Y.Q., Guo, B., Shum, H.Y.: Real-time texture synthesis by patch-based sampling. ACM Trans. Graph. 20(3), 127–150 (2001)
[22] Lefebvre, S., Hoppe, H.: Parallel controllable texture synthesis. ACM Trans. Graph. 24(3), 777–786 (2005)
[23] Lefebvre, S., Hoppe, H.: Appearance-space texture synthesis. ACM Trans. Graph. 25(3), 541–548 (2006)
[24] Barrett, W.A., Cheney, A.S.: Object-based image editing. In: SIGGRAPH 2002: Proceedings of the 29th annual conference on Computer graphics and interactive techniques, pp. 777–784. ACM, New York (2002)
[25] Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(11), 1222–1239 (2001)
[26] Pérez, P., Gangnet, M., Blake, A.: Poisson image editing. ACM Trans. Graph. 22(3), 313–318 (2003)

[27] Rother, C., Kolmogorov, V., Blake, A.: "grabcut": interactive foreground extraction using iterated graph cuts. ACM Trans. Graph. 23(3), 309–314 (2004)
[28] Li, Y., Sun, J., Tang, C.K., Shum, H.Y.: Lazy snapping. ACM Trans. Graph. 23(3), 303–308 (2004)
[29] Agarwala, A., Dontcheva, M., Agrawala, M., Drucker, S., Colburn, A., Curless, B., Salesin, D., Cohen, M.: Interactive digital photomontage. In: SIGGRAPH 2004: ACM SIGGRAPH 2004 Papers, pp. 294–302. ACM, New York (2004)
[30] Jia, J., Sun, J., Tang, C.K., Shum, H.Y.: Drag-and-drop pasting. ACM Trans. Graph. 25(3), 631–637 (2006)
[31] Lalonde, J.F., Hoiem, D., Efros, A.A., Rother, C., Winn, J., Criminisi, A.: Photo clip art. ACM Trans. Graph. 26(3), 3 (2007)
[32] Holland, J.H.: Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor (1975)
[33] Koza, J.R.: Survey of genetic algorithms and genetic programming. In: Proceedings of 1995 WESCON Conference, pp. 589–594. IEEE, Los Alamitos (1995)
[34] Pan, J., McInnes, F., Jack, M.: Application of parallel genetic algorithm and property of multipleglobal optima to vq codevector index assignment for noisy channels. Electronics Letters 32(4), 296–297 (1996)
[35] Liu, B., Liu, B.: Theory and Practice of Uncertain Programming. Physica-Verlag (2002)
[36] Sun, H., Lam, K.Y., Chung, S.L., Dong, W., Gu, M., Sun, J.: Efficient vector quantization using genetic algorithm. Neural Comput. Appl. 14(3), 203–211 (2005)
[37] de Castro, L.N., Zuben, F.J.V.: The clonal selection algorithm with engineering applications. In: Proceedings of GECCO 2000: Workshop on Artificial Immune Systems and Their Applications, Las Vegas, Nevada, USA, August 2000, pp. 36–39 (2000)
[38] de Castro, L.N., Timmis, J.: Artificial Immune Systems: A New Computational Intelligence Approach. Springer, Heidelberg (2002)
[39] de Fran¸ca, F.O., Zuben, F.J.V., de Castro, L.N.: An artificial immune network for multimodal function optimization on dynamic environments. In: GECCO 2005: Proceedings of the 2005 conference on Genetic and evolutionary computation, pp. 289–296. ACM Press, New York (2005)

Efficient Method for Point-Based Rendering on GPUs

Lamei Yan[1] and Youwei Yuan[2]

[1] School of Printing Engineering, Hangzhou Dianzi University
Hangzhou, 310018, China
[2] School of Computer & Software, Hangzhou Dianzi University
Hangzhou, 310018, China
y.lm@163.com

Abstract. We describe methods for high-performance and high-quality rendering of point models, including advanced shading, anti-aliasing, and transparency. we keep the rendering quality as the previous GPU-based point rendering approaches, while involving normal vector computation for each frame. We also present efficient data structures for hierarchical rendering on modern graphics processors (GPUs). In addition we will address methods for geometric processing, filtering and resampling of point models. Examples are presented to illustrate the quality of the meshes produced, and the flexibilities of the computational system.

Keywords: GPUs; mesh simplification; animation; real-time rendering.

1 Introduction

Physically based modeling and animation of deformable objects has been an ongoing research topic in the field of computer graphics since the 80's. We believe that although great achievements have been made, the full potential of alternative modeling primitives and also the possible areas of application have not yet been fully explored. In this research project, we investigate point primitives for modeling and rendering deformable objects, which simplify the simulation of topological changes due to, fracture, melting, clay-like splitting and merging and other possible phenomena.

In the last years point-based rendering has been shown to offer the potential to outperform traditional triangle based rendering both in speed and visual quality when it comes to processing highly complex models. Existing surface splitting techniques achieve superior visual quality by proper filtering but they are still limited in rendering speed.

When considering point-based surface representations in general, we further distinguish between a piecewise constant point sampling [1，2] and piecewise linear surface splats [3]. In this paper we focus on surface splats, since besides providing a higher approximation order, they also allow for more efficient rendering and achieve a higher visual quality by sophisticated anti-aliasing techniques [4].

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 687–694, 2008.

In our paper, we demonstrate how interactive rendering with complex materials can nonetheless be achieved. Rendering of arbitrary materials using this approximation is very fast because it boils down to computing texture coordinates and blending two texture maps together.

The main advantages are:

- Speed: Real-time processing is our major constraint.
- Simplicity: Easy to implement and adapted to further hardware optimizations.
- Smoothness: The visualized surface looks smooth.
- Globally adaptive sampling: Only areas that need accurate sampling are refined.
- Fastest and highest quality: The availability of multiple render targets in combination with a floating point precision rendering pipeline enabled us to derive one of the fastest and highest quality GPU-based surface splatting technique available to date.

2 Related Work

Most of the work in direct volume visualization in recent years has been focused on texture-based approaches. Cabral B[5] and Stegmaier S [6] presented a flexible framework for single pass GPU-raycasting that takes advantage of the easily extensible raycasting approach to demonstrate a number of non-standard volume rendering techniques, including translucent material and self-shadowing isosurfaces, as well as an acceleration technique based on exploiting inter-frame coherence. Hadwiger et al [7] presented a GPU-raycasting system for isosurface rendering. They employ a two-level hierarchical representation of the volume data set for efficient empty space skipping and use an adaptive sampling approach for iterative refinement of the isosurface intersections.

To improve robustness, Ohtake and Belyaev advocated moving the triangle centroids to the zero iso-contour instead of the nodes, and matching the triangle normals with the implicit surface normals [8].

3 Optimal Control Algorithms

To solve the optimal problem, a gradient algorithm with projection, based on instantaneous evaluation of the gradient [9], The diagram of point-based Animation and Real-time Rendering is shown in fig.1

3.1 Point-Cloud Refinement

The main source for real-time point-cloud refinement methods is the work [9] by G. Guennebaud. The rendering of point splats involves several sub-tasks: first the size and shape of the splats have to be determined from the current viewing parameters so that we get a holefree image[10] .Using these techniques alone already results in mid-quality elliptical but still unfiltered surface splats. Nevertheless it provides a much

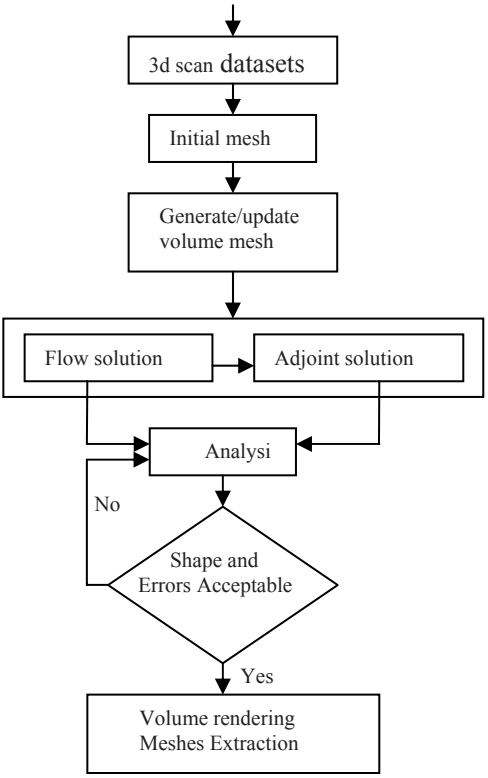


Fig. 1. The diagram of point-based Animation and Real-time Rendering

better representation of the geometry than fixed splat shapes, especially noticeable near contours. Each point in the point set p, have a position, a normal and a radius. The refined point set is iteratively defined. Every weighted point gives rise to a distance function, namely the power distance function,

$$\pi_p : R^3 \to R, x \mapsto \|x - z\|^2 - r \tag{1}$$

Let P be a set of weighted point in R3.The power diagram of P is a decomposition of R3 into the power cells of the points in P[11].

Theorem 1. The natural coordinates satisfy the requirements of a coordinate system, namely, for any p∈P, $\lambda_P(q) = \delta_{pq}$ where δ_{pq} is the Kronecker symbol and the point x is the weighted center of mass of its neighbors. That is,

$$x = \sum_{p \in P} \lambda_p(x) p, with \sum_{p \in P} \lambda_p(x) = 1 \tag{2}$$

Induced distance function.

3.2 Point Based Surface Modeling

Using point-based methods to solve a visualization problem may seem like the wrong way to go since polygonal methods is well known and understood. The graphics hardware is also designed and highly optimized for polygons. We propose the use of point primitives in the context of 3D shape modeling. We split and merge to dynamically adapt the surface sampling density during simulation. To maintain a close connection between physical particles and surface samples, we use a space warping approach, similar to the free-form shape deformation scheme[12] .We use a linear version of the Moving Least Squares projection for dynamic surface reconstruction.

We define a point-based object $P_i^0 = P_i, i = 1,...,n$ as a cloud of surfels and phyxels. where k≥1 [14].

$$P_{2i-1}^k = \omega_1(k)P_i^{k-1} + \omega_0(k)P_{i+1}^{k-1} \ , \tag{3}$$

$$P_{2i}^k = \omega_0(k)P_i^{k-1} + \omega_1(k)P_{i+1}^{k-1} \ ; \tag{4}$$

where, $\omega_0(k) = \dfrac{1}{2(1+\cos(l/2^k))}$, $\omega_1(k) = \dfrac{1+2\cos(l/2^k)}{2(1+\cos(l/2^k))}$, $\omega_0(k)+\omega_1(k)=1$.

3.3 Isosurface Refinement

The point set (point-cloud) P^0, serves as a starting point for the algorithm. The solution is iterative, and for each iteration, points in *Pn* is either accepted as part of the final point-set S, or refined and added to *Pn+1*, which serves as input to the next iteration. Each point $P \in P^n$ is locally refined in each iteration[13]. This procedure contains several steps.

First, the particle p is moved closer to the real isosurface with help of an approximation in form of one Newton-Raphson step

$$p_{n+1} = \mathrm{p_n} - \frac{f(\mathrm{p_n})\nabla f(\mathrm{p_n})}{|\nabla f(\mathrm{p_n})|^2} \tag{5}$$

Where

$$f(\mathrm{p}) = c_S(\mathrm{p}) - k_{isocolor} \tag{6}$$

is defined from gradient of the colour fields[16].

This iterative algorithm continues until the input point set *Pn* is empty or a fixed time-limit is reached.

3.4 Transformation to World Space and Rendering

At this stage, we have a triangle mesh with the correct connectivity and vertices in screen space, e.g. with coordinates $[xp, yp, zp]^T$. In order to render this mesh and to compute reflections and refractions of the 3D environment, the vertices are projected back into world space while the connectivity is kept fixed. We, therefore, need to invert the transformation given in Eq. (1) and Eq. (2). Let $Q \in R^{4\times4}$ be the inverse of

the projection matrix, i.e. $Q = P-1$. With Q, the world coordinates $[xp,yp, zp]^T$ can be computed via the inverse projection equation.

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = Q \begin{bmatrix} (-1+2x_p/W)w \\ (-1+2y_p/W)w \\ z_p \\ w \end{bmatrix} \tag{7}$$

At this point, we do not have the projective divisor w. The GPU stores internally a transformed version of the true depth to achieve better accuracy for near objects. The exact depth can be computed as $d = z_a/(z - z_b)$, where z is the transformed depth from the GPU

$z_a = \dfrac{d_{far}}{d_{far} - d_{near}}$, $z_b = \dfrac{d_{near} d_{far}}{d_{near} - d_{far}}$ and d_{near} and d_{far} are the distances of the near and far clipping planes, respectively. To maximize depth accuracy, the near and far planes are set tightly around the interval of possible depths. The near plane is set slightly closer as the closest point-to-triangle distance of the object vertices, and the far plane is set slightly further away than the maximum distance between cage vertices[14].

4 Numerical Results on Control

In this section, we present an example of how to demonstrate how to capture an analytically defined metric tensor. At each iteration, the discrete analytical metric field is computed at the mesh vertices and a unit mesh is generated with respect to this field. The analytical function is not explicitly used when inserting or moving a vertex. During the re-meshing stage, a linear interpolation scheme is used to find the value of the field at a given vertex location.[15].

In this section, we present an example of mesh adaptation to demonstrate how to capture an analytically defined metric tensor.

In each case, the adapted mesh fits well the metric. The mesh adaptation algorithm statistics indicate that almost 91% of the edges have a unit length, i.e., a length between $1/\sqrt{2}$ and $\sqrt{2}$ .We can compute the efficiency index of the resulting adapted meshes, i.e., a scalar value representing the adequacy between the metric specification and the actual element size, with the following formula:

$$T_H = \exp\left(\frac{1}{n_e} \sum_{i=1}^{n_e} (Q_l(e_i) - 1) \right) \tag{8}$$

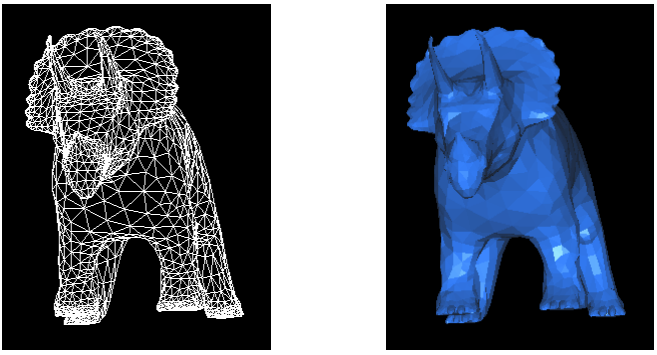
Where n_e is the number of edges of the mesh and $Q_l(e_i)$ is the length quality of the edge e_i in the metric given by:

$$Q_l(e_i) = \begin{cases} l_M(e_i) & if \quad l_M(e_i) \le 1 \\ (l_M(e_i))^{-1} & \text{else} \end{cases} \tag{9}$$

With $l_M(e_i)$ the edge length in the metric M. Here, an efficiency index close to 0.86 is obtained in each case. $d = \sqrt{x^2 + y^2 + z^2}$

$$f_{3d}(x, y, z) = \begin{cases} 0.1\sin(50x) & \text{if } x \le \frac{-\pi}{50} \\ \sin(50x) & \text{if } \frac{-\pi}{50} < x \le \frac{2\pi}{50} \\ 0.1\sin(50x) & \text{if } \frac{2\pi}{50} < x \end{cases} \tag{10}$$

See figure 2 for an example rendered with this technique at real-time rates.



(a) (b)

Fig. 2. Triceratops point-based Animation and Real-time Rendering(Vertices number : 4209; triangles numbers:5008; tetrahedral:16910) (a) mesh generation from point-sets (b) Hardware accelerated (>60Hz) rendering.

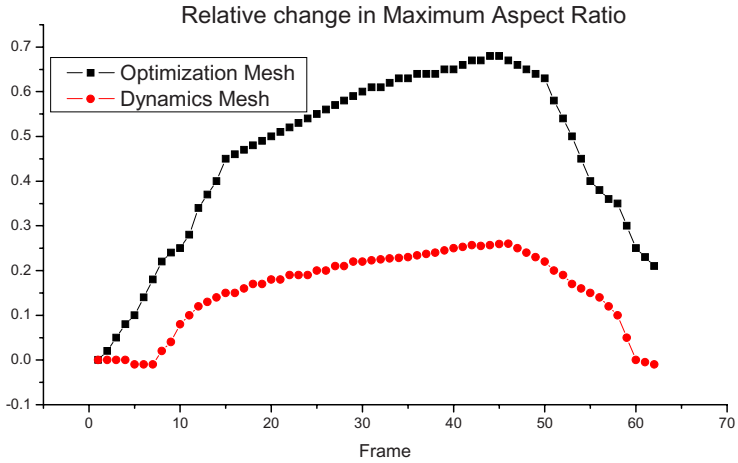


Fig. 3. Plot of changes in maximum aspect ratios during simulation of isometric contraction for dynamics and optimization based meshes.

Figure 3 shows the relative change in maximum aspect ratio observed during an isometric contraction of the biceps for meshes created using the optimization algorithm and using the dynamics algorithm. Similar results were observed for the triceps and during isotonic contraction [16]. These results suggest that initial mesh quality may be misleading and not sufficient to guarantee performance of a mesh throughout simulation. In all of our comparisons, the optimization based meshes were of higher quality initially, but tended to undergo as much as a 67% change in maximum aspect ratio during muscle contraction, whereas the dynamics based meshes tended to degrade by only 20.02%.

5 Conclusions and Future Works

We presented an algorithm for producing a high quality unstructured mesh directly from the input datasets. We also provides an overview of methods that can be used for a point-based rendering pipeline, from isosurface extraction to surface shading. The metric used for 3D simulations is constructed with an a posteriori error estimate based on a discrete approximation of Hessian of the solution.

Our approach fully demonstrates it's ability for noisy point sets by the noise reduction process. Normal vectors are computed and various types of noises are reduced on such an image buffer.

The methods are compared with respect to ease of implementation, performance and shading quality. With the introduced tools, visualization of unstructured meshes can be performed in real-time, on the latest graphics hardware.

The main identified problem is the surface refinement step. A real-time point-set surface refinement method, that work along the simulation and rendering, is yet to be found. In future work, we wish to decrease flickering effects in interactive rendering, and to apply our approach for irregular point sets.

References

1. Pfister, H., Zwicker, M., Van Baar, J., Gross, M.: Surfels: Surface Elements as Rendering Primitives. In: Proc. of ACM SIGGRAPH 2000, pp. 335–342 (2000)
2. Alexa, M., Behr, J., Cohenor, D., Fleishman, S., Levin, D., Silva, C.T.: Point Set Surfaces. In: Proc. of IEEE Visualization 2001, pp. 21–28 (2001)
3. Boulic, R., Magnenat-Thalmann, N., Thalmann, D.: A Global Human Walking Model with Real Time Kinematic Personification. The Visual Computer 6, 344–358 (1990)
4. Kobbelt, L., Botsch, M.: A Survey of Point-based Techniques in Computer Graphics. Computers & Graphics 28(6), 801–814 (2004)
5. Cabral, B., Cam, N., Foran, J.: Accelerated Volume Rendering and Tomographic Reconstruction using Texture Mapping Hardware. In: Proceedings of the 1994 Symposium on Volume Visualization, pp. 91–98 (1994)
6. Stegmaier, S., Strengert, M., Klein, T., Ertl, T.: A Simple and Flexible Volume Rendering Framework for Graphics-Hardware based Raycasting. In: Proceedings of the International Workshop on Volume Graphics 2005, pp. 187–195 (2005)
7. Hadwiger, M., Sigg, C., Scharsach, H., Bhler, K., Gross, M.: Real-Time Ray-Casting and Advanced Shading of Discrete Isosurfaces. In: Proceedings of Eurographics 2005, pp. 303–312 (2005)

8. Ohtake, Y., Belyaev, A.G.: Dual/Primal Mesh Optimization for Polygonized Implicit Surfaces. In: Proc. of the 7th ACM Symp. on Solid Model. and Appl., pp. 171–178. ACM Press, New York (2002)
9. Guennebaud, G., Barthe, L., Paulin, M.: Interpolatory Refinement for Real-Time Processing of Point-Based Geometry. In: Computer Graphics Forum, Eurographics 2005 conference proceedings, September 2005, vol. 24, pp. 657–667 (2005)
10. Frey, P., Alauzet, F.: Anisotropic Mesh Adaptation for CFD Computations. Comput. Methods Appl. Mech. Engrg. 194, 48–49 (2005)
11. Guennebaud, G., Barthe, L., Paulin, M.: Dynamic Surfel Set Refinement for High Quality Rendering. In: Computer & Graphics, pp. 827–838. Elsevier Science, Amsterdam (2004)
12. Alexa, M., Adamson, A.: On Normals and Projection Operators for Surfaces Defined by Point Sets. In: Proceedings of Eurographics Symposium on Point-Based Graphics 2004, pp. 150–155 (2004)
13. Frey, P.J.: About Surface Remeshing. In: Proc.of 9th Int. Meshing Roundtable, New Orleans, LO, USA, pp. 123–136 (2000)
14. Ganovelli, F., Cignoni, P., Montani, C., Scopigno, R.: A Multiresolution Model for Soft Objects Supporting Interactive Cuts and Lacerations. In: Eurographics, pp. 271–282 (2000)
15. You-wei, Y., La-mei, Y.: A Neural Network Approach for 3-D Face Shape Reconstruction. In: Proceedings of 2002 International Conference on Machine Learning and Cybernetics, vol. 5, pp. 2073–2077. IEEE, Beijing (2002)
16. Bordeux, C., Boulic, R., Thalmann, D.: An Efficient and Flexible Perception Pipeline for Autonomous Agents. In: Proceedings of Eurographics 1999, Milano, Italy, pp. 23–30 (1999)

Efficient Mushroom Cloud Simulation on GPU

Xingquan Cai, Jinhong Li, and Zhitong Su

College of Information Engineering, North China University of Technology,
Beijing, 100041, China
xingquancai@126.com

Abstract. In this paper, we present a method to simulate the Mushroom Cloud efficient on GPU using advanced particle system, and our particle system is a state-preserving simulation system. We provide the visual-only model of Mushroom Cloud and we divide the Mushroom Cloud into five portions. Then we present our advanced particle system method. Our particle system method processes the birth and death of particles via index on CPU and uses a pair of Floating Point Textures on GPU to store the dynamic attributes of particles. This method also updates the dynamic attributes of particles and renders the system on GPU. We also provide a three-layers hierarchical structure to manage the particle system and batch rendering the particles having the similar attributes. Finally, the experiments prove that our method is feasible and high performance.

Keywords: Mushroom Cloud simulation, particle system, GPU (Graphics Processing Unit), Floating Point Textures, state-preserving simulation.

1 Introduction

Most natural sceneries effects, such as clouds, fire, rain, snow, smoke, sparks, blood, etc, are full of motion, full of chaos and full of fuzzy objects, and change with time past. These natural sceneries effects are routinely created in todays video games. So the simulation of natural sceneries becomes a hot topic in the research field of computer graphics.

Usually, the particle systems could be used to simulate these natural sceneries on CPU [1, 2, 3, 4, 5]. However, if the number of particles is above 10K, the particle system on CPU is difficult to run on real-time. It is required that plenty of particles more than 10K should be need in simulate system of photorealist natural sceneries effects. Today, with the development of GPU, we could deal with complex computing and programming on GPUs.

In this paper, we present a new method to simulate the Mushroom Cloud efficient on GPU using advanced particle system, and our particle system is a state-preserving simulation system. We provide the visual-only model of Mushroom Cloud and we divide the Mushroom Cloud into five portions. Then we present our advanced particle system method in detail. Our particle system method processes the birth and death of particles via index on CPU and uses a

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 695–706, 2008.

pair of Floating Point Textures on GPU to store the dynamic attributes of particles. This method also updates the dynamic attributes of particles, handles the collision between particles and other models and renders the system on GPU. We also provide a three-layers hierarchical structure to manage the particle system and batch rendering the particles having the similar attributes.

In this paper, after exploring the related work of particle system, we present our visual-only model of Mushroom Cloud. Then we introduce our particle system method on GPU. In Section 5, we give the three-layers hierarchical structure of our advanced particle system. In Section 6, we show the results using our method before we draw the conclusion in Section 7.

2 Related Work

Particle system has a long history in video games and computer graphics. Very early video games in the 1960s already used 2D pixel clouds to simulate explosions. In 1983, Reeves [6] first described the basic motion operations and the basic data representing a particle, which both have not been altered much since being presented. In 1990, an implementation on parallel processors of a super computer has been done by Sims [7]. In 2000, McAllister [8] also described many of the velocity and position operations of the motion simulation which are used in this paper. The latest descriptions of CPU-based particle system for using in video games and photorealist natural sceneries have been done by Wang et al [1], Liu et al [4], Burg [9].

With the development of GPU, several forms of physical simulation have recently been developed for modern GPU. In 2003, Harris [10] has used GPU to perform fluid simulations and cellular automata with similar texture-based iterative computation. Green [11] describes a cloth simulation using simple grid-aligned particle physics, but does not discuss generic particle systems' problems, like allocation, rendering and deallocation. However, their algorithms do not show the necessary properties to exploit the high frame-to-frame coherence of the particle system simulation. Recently, Schneider et al [12], Li et al [13], Livny et al [14], and Eric et al [15] have used GPU to render large scale terrain scene. Christopher et al [16] also provide the method of real-time mesh simplification using GPU. As GPU could deal with complex computing so fast, we want to implement particle system on GPU.

Some particle systems have been implemented with vertex shaders (also called vertex programs) on programmable GPUs in NVIDIA SDK [17]. However, these particle systems are stateless. They do not store the current attributes of the particles, including current position, current velocity, etc. To determine a particle's position, the system needs to find a closed form function for computing the current position only from initial values and the current time. Stateless particles are not meant to collide with the environment. They are only influenced by global gravity acceleration and could be simulated quite easily with a simple function. As a consequence, such particle system can hardly react to a dynamic

environment. However, rather complex effects usually need collisions or forces with local influence.

The strengths of the stateless particle system make it ideal for simulating small and simple effects without influence from the local environment. But in action video games, these might be a weapon impact splash or the sparks of a collision. Stateless particles method is less suitable for these larger effects that require interaction with the environment. So we provide a state-preserving particle system method in this paper. Our method is very similar to Lattas method [18] and Kolbs method [19], but our method is simpler than them.

There are a few methods to simulate the Mushroom Cloud, so in this paper we provide an efficient method to simulate the Mushroom Cloud. Firstly, we provide our visual-only model of Mushroom Cloud.

3 Visual-Only Model of Mushroom Cloud

In this paper, we give up the physical theory of Nuclear Explosions, and we provide the visual-only model of Mushroom Cloud. After making a systematic observation of the photo and video of Mushroom Cloud, just as Fig. 1 shows, we divide the Mushroom Cloud into five portions, including Bottom Wave portion, Ground Shock Wave portion, Column portion, Ring portion and Core portion.

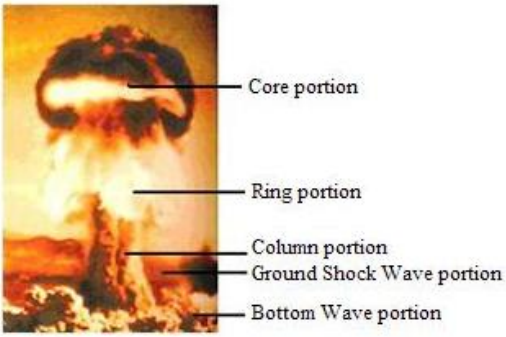


Fig. 1. The Mushroom Cloud

3.1 Bottom Wave Portion

The Bottom Wave portion is like a cake and fills in the blank between the Ground Shock Wave portion and the Core portion. The thickness of the cake is not equable, and the part near the core is thicker. The thickness could be described used the function, $y = k/x$. Firstly, the Bottom Wave portion is expanding. Then, it is shrink. In Fig. 2, a) is the cutaway view, and b), c) are the planforms. a) shows that the center of the Bottom Wave portion is rising and the margin is expanding. b) is the expanding process and c) is the shrink process.

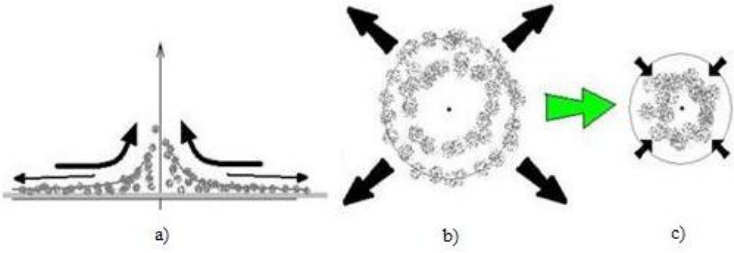


Fig. 2. Bottom Wave portion

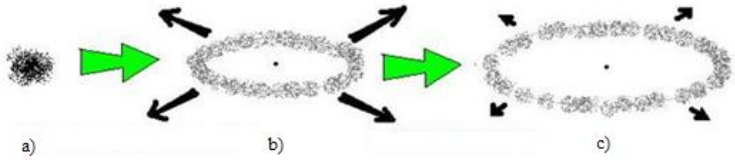


Fig. 3. Ground Shock Wave portion

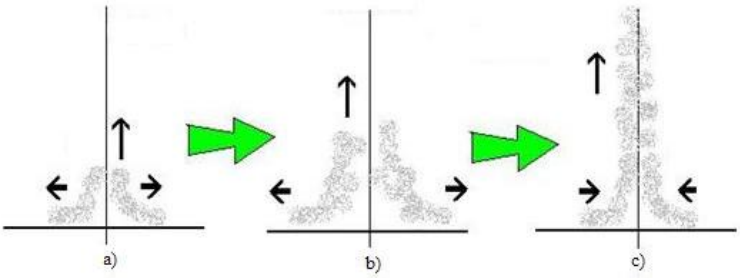


Fig. 4. Column portion

3.2 Ground Shock Wave Portion

When the nuclear bomb is exploding, the Ground Shock Wave sticks to the terrain. And the core of Ground Shock Wave expands to a bigger ring of fragments and smoke fast. The Ground Shock Wave is not expanding in uniform speed. At the beginning, Ground Shock Wave is expanding fast. Then, the bigger ring of fragments and smoke is expanding in reduced speed. Finally, Ground Shock Wave becomes transparency and fades away. Fig. 3 shows this process. a) is expanding to b) in a fast speed. And b) is changing to c) in slowdown speed.

3.3 Column Portion

The Column portion is the smoke column of Mushroom Cloud from the terrain surface to the air. It's shape like a curved column, so we use an inverse proportion function to create it. The Column portion joints the Core portion and the Bottom Wave portion. In Fig. 4, as the Core portion is rising, the height of Column portion becomes higher and the radius is outspread. Then the radius is shrinking.

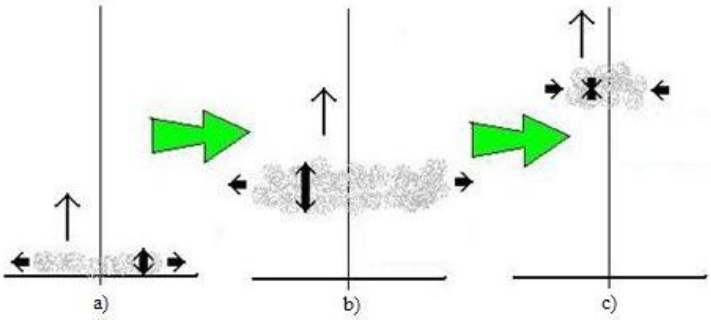


Fig. 5. Ring portion

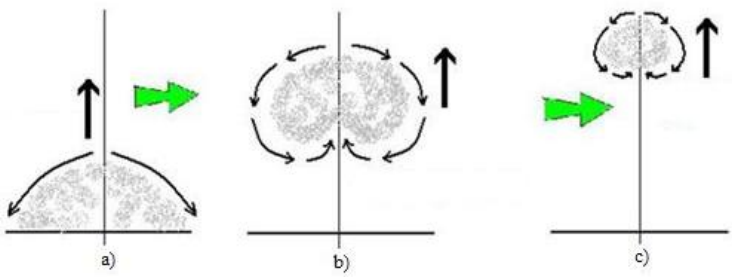


Fig. 6. Core portion

3.4 Ring Portion

Usually, the Ring portion is coming into being in the bigger explosion equivalent of nuclear munitions, and is hackneyed in H-Bomb Explosions and Neutron Bomb Explosions. It is an expanding ring similar to Ground Shock Wave portion. But its moving up as time past. The Ring portion is coming into being up the terrain surface, and is rising with the radius expanding. And at some poison, the radius is shrinking. Fig. 5 shows this process.

3.5 Core Portion

At the beginning of explosion, the Core portion is not the mushroom shape. It is close to a hemisphere. And its position is very low. As the time past, the hemisphere shape Core portion is moving up and changing to the Mushroom Cloud that we all know. We use the hemi-ellipse function and the Double Folium function to create the Core portion. As the Core portion rising, the Mushroom Cloud is shrinking. Fig. 6 shows the changing process.

4 Particle System on GPU

The following subsections describe the algorithm of our state-preserving particle system on GPU in detail.

The algorithm consists of four basic steps:

1. Processing birth and death
2. Updating particles attributes with collision detection
3. Transferring texture data to vertex data
4. Rendering particles

4.1 Particle Data Storage

Position is one of the most important attributes of a particle. In our system, positions of all active particles are stored in a floating point texture with three color components that will be treated as x, y and z coordinates. Each texture is conceptually treated as a one-dimensional array, texture coordinates representing the array index. However, the actual textures need to be two-dimensional for the size restrictions of current hardware. The texture itself is also a render target, so it can be updated with the computed positions. In the stream processing model [20], which is the programming model in graphics hardware, it represents either the input or the output data stream. As a texture cannot be used as input and output at the same time, we use a pair of these textures and a double buffering technique to compute new data from the previous values.

If other particle attributes, such as velocity, orientation, size, color, and opacity, were to be simulated with the iterative integration method, they would need texture double buffers as well. However, since these attributes typically follow simple computation rules or are even static, we can take a simpler approach, and we just take the important attributes like position, velocity, color, etc. And other static attributes just need one texture buffer.

4.2 Particle Birth and Death

The particles in a system can either exist permanently or only for a limited time. A static number of permanently existing particles represents the simplest case for the simulation, because it only requires uploading all initial particle data to the particle attributes textures once. As this case is rather rare, we assume

a varying number of short-living particles for the rest of the discussion. The particle system must then process the birth of a new particle like its allocation, the death of a particle and its deallocation.

The birth of a particle requires associating new data with an available index in the attribute textures. Since allocation problems are serial by nature, this can not be done efficiently with a data-parallel algorithm on the GPU. Therefore an available index is determined on the CPU via traditional fast allocation schemes. The simplest allocation method uses a stack filled with all available indices.

In our method, the particle's death is processed independently on the CPU and GPU. The CPU registers the death of a particle and adds the freed index to the allocator. The GPU does an extra pass over the particle data: The death of a particle is determined by the time of birth and the computed age. The dead particle's position is simply moved to invisible areas, e.g. infinity. As particles usually fade out or fall out of visible areas anyway at the end of their lifetime, the extra pass rarely really needs to be done. It is a basically clean-up step to increase rendering efficiency.

4.3 Update Particles Attributes

The most important attributes of a particle are its position and velocity. So we just deal with particle position and velocity. The actual program code for the attributes simulation is a pixel shader which is used with the stream processing algorithm. The shader is executed for each pixel of the render target by rendering a screen-sized quad. The current render target is set to one of the double buffer attribute textures. The other texture of the double buffer is used as input data stream and contains the attributes from the previous time step. Other particle data, either from inside the attribute textures or as general constants, is set before the shader is executed.

Update Velocities: There are several velocity operations that can be combined as desired: global forces (e.g. gravity, wind), local forces (attraction, repulsion), velocity dampening, and collision responses. For our GPU-based particle system these operations need to be parameterized via pixel shader constants. Their dynamic combination is a typical problem of real-time graphics. Comparable to the problem of light sources and material combinations, it could be solved in similar ways. Typical operation combinations are to be prepared in several variations beforehand. Other operations could be applied in separate passes, as all operations are completely independent.

Global and local forces are accumulated into a single force vector. The acceleration can then be calculated with Newtonian physics as Formula 1. In Formula 1, a is the acceleration vector, F is the accumulated force and m is the mass of the particle. If all particles have unit mass and forces have the same value, the accelerations can be used without further computation.

$$a = \frac{F}{m}. \tag{1}$$

$$v = \overline{v} + a \cdot \triangle t. \tag{2}$$

The velocity is then updated from the acceleration with a simple Euler integration in the form of Formula 2. In Formula 2, v is the current velocity, $\overline{v}$ is the previous velocity and $\triangle t$ is the time step.

$$p = \overline{p} + v \cdot \triangle t. \tag{3}$$

Update Positions: Euler integration has already been used in the previous section to update the velocity by using the acceleration. The computed velocity can be applied to all particles in just the same way. We use Formula 3 to update the position. In Formula 3, p is the current position and $\overline{p}$ is the previous position.

4.4 Transfer Texture Data to Vertex Data

Before rendering particles, we should copy the particle data from the floating point texture to vertex data. Copying the particles data from a texture to vertex data is a hardware feature that is only just coming up in PC GPUs. OpenGL [21] offers vertex textures with ARB_vertex_shader extension. OpenGL also provides two functions, glReadBuffer and glReadPixels. These functions could copy the particle data from the floating point texture to vertex data.

4.5 Render Particles

The particles can be rendered as point sprites, triangles or quads. If the particle is rendered as triangles or quads, the particle may have three vertices or more vertices. So it is required that we should recompute the vertices position of particle before rendering. To avoid this overhead, our implementation uses point sprites.

5 Three-Layers Hierarchical Structure Particle System

The impressive improvement of graphics hardware in terms of computation and communication speed is reshaping the real-time rendering domain. A number of performance and architectural aspects have a major impact on the design of real-time rendering methods. Todays GPUs are able to sustain speeds of hundreds of millions of triangles per second and can render triangles in batches. Usually, in video games and photorealist natural sceneries, there are plenty of particles. In order to use the ability of GPUs rendering triangles in batches and be convenient to manage the particle system, we import the theory of advanced particle system [9].

Just as Fig. 7 shows, we divide the particle system into three layers, including Particles Manager, Particles Cluster and Particles. Particles Cluster is the batches of particle having the similar attributes, such as velocity, color, texture, etc. Particles Manager manages the Particles Clusters and is responsible

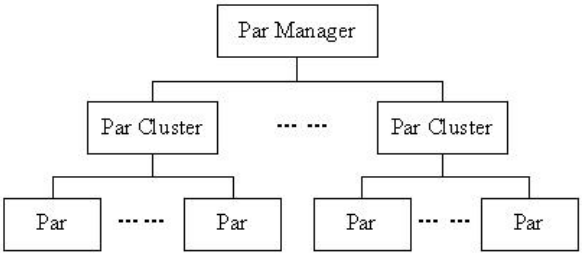


Fig. 7. Three-layers hierarchical structure particle system ('Par' stands for 'Particle')

for the birth of a new Particles Cluster, the death of a Particles Cluster and its deallocation. In this way, we could use the ability of GPUs rendering triangles in batches and implement plenty of particles in video games and photorealist natural sceneries.

6 Results

We have implemented our algorithm. Our implementations are running on a Intel PIV 2.8GHz computer with 1GB RAM, and NVIDIA GeForce7650 graphics card with 256M RAM, under Windows XP, Visual C++6.0, OpenGL and Cg environment, while running smoothly in real time. The rendering system has a real viewport size of 1024×768 .

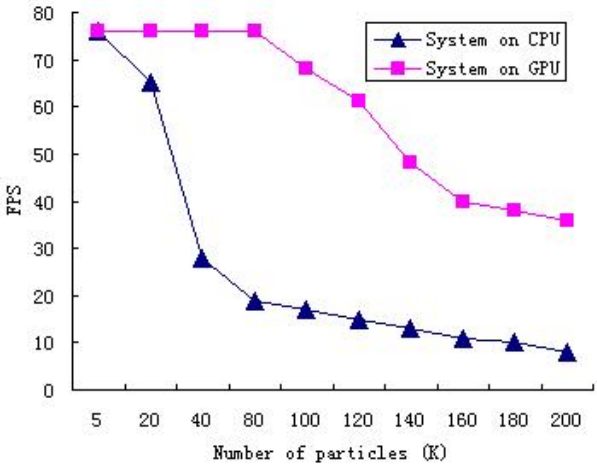


Fig. 8. Comparison between particle system on CPU and particle system on GPU

6.1 Comparison with CPU Particles

We implement a particle system on CPU and a particle system on GPU to simulate flowing magma. There is only one Particle Cluster in each system. The particle just has the gravity and we do not concern other forces and the collision. At the same number of particles, we note the rendering frame rate. In order to ensure the objectivity of the experiment data, we sample continuous 1000 frames, note the FPS (Frames Per Second) and compute the average FPS.

Just as Fig. 8 shows, in our experiment, when the number of particles is 100,000, the FPS of particle system on GPU is above 60. But at the similar condition, the FPS of particle system on CPU is below 18. When the number of particles is 200,000, the FPS of particle system on GPU is 36, but the FPS of particle system on CPU is below 8. All these prove that particle system on GPU is higher performance than particle system on CPU.

6.2 Complex Mushroom Cloud Effects

We also have used our method to simulate the Mushroom Cloud of atomic bomb explosion effect. The mushroom cloud has five Particles Clusters. Five Particles Clusters stand for Bottom Wave portion, Ground Shock Wave portion, Column portion, Ring portion and Core portion of the mushroom cloud. All the particles of five Particles Clusters are above 60,000, and the system is running smoothly and is above 25 fps. Fig. 9 shows the process of Mushroom Cloud effects.

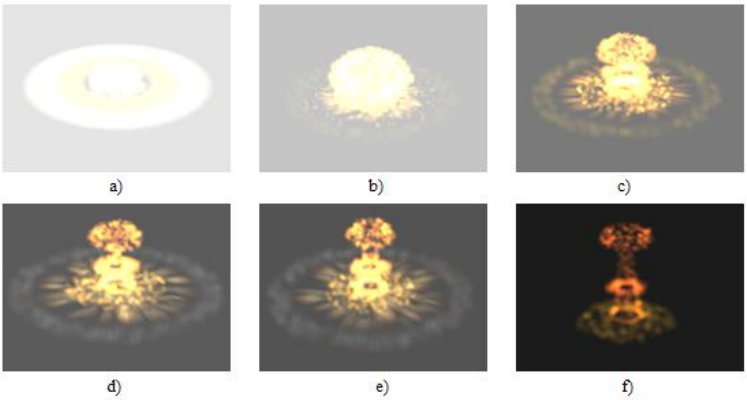


Fig. 9. The Mushroom Cloud of atomic bomb explosion

7 Conclusion and Future Work

In this paper, we present a new method to simulate the Mushroom Cloud efficient on GPU using advanced particle system, and our particle system is a state-preserving simulation system. We provide the visual-only model of Mushroom

Cloud and we divide the Mushroom Cloud into five portions. Then we present our advanced particle system method in detail. Our particle system method processes the birth and death of particles via index on CPU and uses a pair of Floating Point Textures on GPU to store the dynamic attributes of particles. This method also updates the dynamic attributes of particles, handles the collision between particles and other models and renders the system on GPU. We also provide a three-layers hierarchical structure to manage the particle system and batch rendering the particles having the similar attributes. Finally, the experiments prove that our method is feasible and high performance.

As a future possibility, we are working on using our method to implement other complex natural phenomenon, and developing the method to deal with the collision between particles and other particles.

Acknowledgments. This work was supported by PHR(IHLB) (Funding Project for Academic Human Resources Development in Institutions of Higher Learning under the Jurisdiction of Beijing Municipality) Grant. We would like to thank those who care of this paper and our projects. Also, we would like to thank everyone who spent time on reading early versions of this paper, including the anonymous reviewers. And thanks to those who devote themselves into studies on Graphics, they gave me inspirations as well as wonderful demos of their works.

References

1. Changbo, W., Zhangye, W., Qunsheng, P.: Real-time Snowing Simulation. The Visual Computer 22(5), 315–323 (2006)
2. Huagen, W., Xiaogang, J., Qunsheng, P.: Physically based real time fountain simulation. Chinese Journal of Computers 21(9), 774–779 (1998)
3. Ruofen, T., Lingjun, C., Guozhao, W.: A method for quick smog simulation. Journal of Software 10(6), 647–651 (1999)
4. Xiaoping, L., Ye, Y., Hao, C., et al.: Real-time simulation of special effects in navigation scene. Journal of Engineering Graphics (3), 44–49 (2007)
5. Yu, G., Lin-can, Z., Wei, C., Qun-sheng, P.: Real Time Waterfall Simulation Based Particle System 16(11), 2471–2474 (2004)
6. Reeves, W.T.: Particle Systems-Technique for Modeling a Class of Fuzzy Objects. In: Proceedings of SIGGRAPH 1983 (1983)
7. Karl, S.: Particle Animation and Rendering Using Data Parallel Computation. Computer Graphics 24(4), 405–413 (1990)
8. McAllister, D.K.: The Design of an API for Particle Systems. Technical Report, Department of Computer Science, University of North Carolina at Chapel Hill (2000)
9. Burg, V.D.: Building an Advanced Particle System. Game Developer Magazine (2000), http://www.gamasutra.com/features/20000623/vanderburg_pfv.htm
10. Harris, M.: Real-Time Cloud Simulation and Rendering. PhD thesis, University of North Carolina at Chapel Hill (2003)
11. Green, S.: Stupid OpenGL Shader Tricks (2003), http://developer.nvidia.com/docs/IO/8230/GDC2003_OpenGLShaderTricks.pdf

12. Schneider, J., Westermann, R.: GPU-Friendly High-Quality Terrain Rendering. Journal of WSCG 14, 49–56 (2006)
13. Sheng, L., Junfeng, J., Xuehui, L., Enhua, W.: High performance navigation of very large-scale terrain environment. Journal of Software 17(3), 535–545 (2006)
14. Livny, Y., Kogan, Z., El-Sana, J.: Seamless Patches for GPU-based Terrain Rendering. In: Proceedings of WSCG 2007, pp. 201–208 (2007)
15. Eric, B., Fabrice, N.: Real-time rendering and editing of vector-based terrains. In: Proceedings of Eurographics 2008 (2008)
16. Christopher, D., Natalya, T.: Real-time Mesh Simplification Using the GPU. In: Proceedings of Symposium on Interactive 3D Graphics 2007 (I3D 2007), p. 6 (2007)
17. NVIDIA Corporation: NVIDIA SDK (2004), http://developer.nvidia.com
18. Latta, L.: Building a Million Particle System. In: Proceedings of Game Developers Conference 2004 (GDC 2004) (2004)
19. Kolb, A., Latta, L., et al.: Hardware-based Simulation and Collision Detection for Large Particle Systems. In: Proceedings of Graphics Hardware 2004, pp. 123–132 (2004)
20. Ian, B.: Data Parallel Computing on Graphics Hardware. Stanford University (2003)
21. OpenGL ARB: OpenGL Extension ARB_vertex_shader (2003), http://oss.sgi.com/projects/ogl-sample/registry/ARB/vertex_shader.txt
22. Cai, X., Li, F., et al.: Research of Dynamic Terrain in Complex Battlefield Environments. In: Pan, Z., Aylett, R.S., Diener, H., Jin, X., Göbel, S., Li, L. (eds.) Edutainment 2006. LNCS, vol. 3942, pp. 903–912. Springer, Heidelberg (2006)
23. Fernando, R., Mark, J.K.: The Cg Tutorial: The Definitive Guide to Programmable Real-Time Graphics. Addison Wesley Publishing, Reading (2003)
24. Fernando, R.: GPU Gems: Programming Techniques, Tips, and Tricks for Real-Time Graphics. Addison Wesley Publishing, Reading (2004)
25. Matt, P.: GPU Gems 2: Programming Techniques for High-Performance Graphics and General-Purpose Computation. Addison Wesley Publishing, Reading (2005)

Virtual Artistic Paper-Cut

Hanwen Guo[1,*], Minyong Shi[2], Zhiguo Hong[1], Rui Yang[1], and Li Zhang[1]

[1] Computer School, Communication University of China, Beijing 100024, China
[2] Animation School, Communication University of China, Beijing 100024, China
Eniak_19@msn.com, myshi9613@hotmail.com

Abstract. This paper presents some algorithms for a novel simulation of the folding and cut crafts in Chinese Paper-cut. All algorithms in simulation are designed by consideration of the real Paper-cut's whole process. We applies area filling algorithm, polygon clipping algorithm, contour extraction algorithm and subdivision curve algorithm, with the proper improvement that meets the technique of the art in the nature, to yield some vivid Paper-cut works. Our approach is feasible to anyone with the desire, due to Paper-cut illumines itself in 2D and all the operations are also in 2D.Moverover, recent compelling graphic systems, such as Flash, Photoshop, Maya and 3DMax, did not provide a similar interface for paper's fold and decorative pattern which are primary elements in Paper-cut. However, our approach not only meets the interest of the Paper-cut, but also holds an interface for these compelling systems because of measuring up the SVG.

Keywords: NPR; Polygon Clipping; Contour Extraction; Subdivision Curve; Paper-Cut; SVG.

1 Introduction

The Chinese civilization can be traced far back in ancient times, with a long history of 5,000 years. As the traditional Chinese painting and Chinese paper-cut branded with "Chinese style", embed by lasting appeal, they are widely recognized and well received by the world arts community and academia. However, because of the high cost and long production cycle, it's unable to meet the development requirements by using the Chinese traditional art methods to produce works of art.

Paper-cut is one of the most popular folk art in China; according to the archaeology, its history can go back to the sixth century AD. As a form of folk art, paper-cut inherits the traditional skills and modeling, and absorbs the experience of skills and techniques accumulated by great people in China, flickering aesthetic concept for working people.

---

[*] Hanwen Guo was born in 1983 in China. Currently, he is a postgraduate student is Computer Graphics in Communication University of China. His specific research area is computational geometry and motion retargeting technique. Professor Minyong Shi, who was born in 1964 in China, he received his PhD degree in Beijing Institute of Technology, his major research area is Graph Theory.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 707–718, 2008.

In recent years, as the development of non-photorealistic rendering technology, some new technologies are being used to the traditional artistic creations, thus, the traditional arts with Chinese characteristics are refreshed lives by current.

3D Max and Maya are the world-renowned powerful animation software and bring the convenience and effect to people. The corporations, such as Disney and Dream-Works, produce splendid productions by using this commercial software. But it's inconvenient to make paper folding and cutting animation by the 3D Max or Maya.

For example, the hollow out in paper, as the ordinary craft, is hard to implement on a 3D modeling. The artists need to do lots of affine transformations with the modeling to keep relations between the decorative patterns are correct. But, in reality, it often comes out the unpleasant or inappropriate results. They are forced to do lots of the annoying work and spend a large time on redesign. It is estimates that the whole efforts spent on a dramatis personae are about a week.

In addition, the craft of paper-cut are mostly mastered by old artists. 3D Max and Maya are so powerful and complex that the old artists always feel comfortable when they are using, and everyman, too. For them, the 2D software is more close to their habits and recall us the true nature of paper-cut skills on some extent. But no reports in 2D compelling graphic system, Photoshop and Flash, appeal this need.

To meet this need currently, many domestic colleges and universities have started to investigate the paper-cut using the non-photorealistic rendering technology. Related work [1] has been published on the domestic core journal. Many related articles mainly concentrated on rendering of material and so on. In the field of paper-cut, some articles about paper-cut's decorative patterns are published, and the research and development of three-dimensional paper-cut software has also made great success in China.

As many traditional paper-cut works are in the two-dimensional, it's becoming an urgent research topics to develop a fast, efficient and convenient paper-cut animation software to fold and cut the "virtual paper" in computer, and it's expected to be a good platform for producing many outstanding paper-cut works, fast to make up for the blank in relevant fields. So far we have done many researches on 2D paper's folding and cutting, and develop the paper-cut animation software with the VS2005.

2 Relative Algorithms and Our Classes and Date Structure

2.1 Relative Algorithms

Three main algorithms, as present in this paper, are Folding Algorithm, Cutting Algorithm and Unfolding algorithm. The Reform Outline algorithm, which considered as incidental, is employed by Unfolding Algorithm. The algorithms we have designed are the improvements of many kinds of algorithms in the Computer Graphics, which are belong to four main research directions, area filling algorithm, polygon cutting algorithm, contour extraction algorithm and curve subdivision algorithm.

In the traditional raster display, polygon area filling is an important research area. The representative filling algorithm based on the scanning line is [3].

The Sutherland[4] algorithm and Weiler-Atherton [5] algorithm are the famous algorithms in the area of polygon clipping. Algorithm [4] is applied to rectangle polygon, while algorithm [5] is applied to any convex and concave polygon.

The contour extraction algorithm can be roughly summed up as the following two types: one is the Snake [6] algorithm and another one is Active Contour Models [7] algorithm. This kind of algorithms first need the initial outline and then iteration so that the contour gets close along the direction of the energy reduction, finally get an optimization border.

Over the past few decades, researchers has proposed a lot of mature curve subdivision algorithms, such as the angle cutting curve subdivision algorithm early proposed by Chaikin [8], de Casteljau algorithm [9], de Boor algorithm [10], four points curve subdivision algorithm [11] proposed by Dyn and other researchers and Hermite subdivision algorithm [12].

In our algorithms, we make some improvement of the polygon filling algorithm base on the scanning line and apply it to the Folding Algorithm. Polygon clipping algorithm, Weiler-Atherton, as the definition of the Cutting Algorithm, is improved to be more suitable for the actual situation in the paper-cut. At the same time, we employ contour extraction algorithm and curves subdivision algorithm for management of the decorative patterns and complex shapes. The Unfolding Algorithm, in some words, is the converse process of Folding Algorithm and the combination of the improved Weiler-Atherton.

2.2 Classes and Data Structure in This Paper

As shown in figure 1, the classes in our algorithms are: Paper, Polygon, FlodPolygon and CutPolygon.

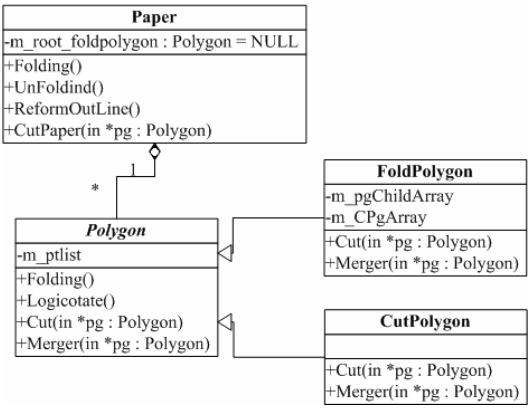


Fig. 1. Relation

Paper is a class equivalent to a container and it has a pointer pointing to the root node of a polygon's multi-tree. Class Paper has many member methods, such as Init (Polygon * pg), Folding (Axis * axis), Unfolding (), and someone else.

In Paper::Init (Polygon * pg), Paper:: m_root_polygon will automatically point to the pointer of polygon, and set itself as the root node of the folding polygon's tree.

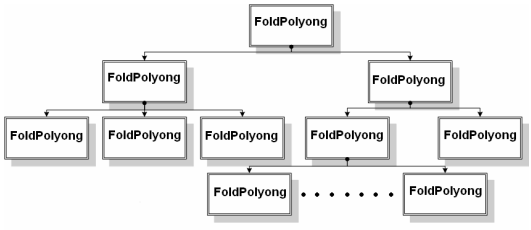


Fig. 2. Mutil-Tree

In Paper::Folding (Axis * axis), a folding axis's pointer is imported in each time for the operation fold between the folding axis and the leaf nodes of the polygons' tree to yield some new nodes. The new nodes are added to the tree. At the same time, the folding axis also be recorded on the current folded nodes when is able to fold, otherwise, it does not.

In Paper::UnFolding(), the patterns generated according to the leaf nodes of the mutil-tree which as the root node pointed by Paper:: m_root_polygon and the initial polygon which is imported by the Init(Polygon* pg) are the objects we will manage. In every cutting, the results of cuttings will not exceed the areas of the primary Fold-Polygon. And in every fold, the information of the folding axis will be recorded on the folded node. Because of the existence of such information, it allows us to modify the shape of the paper easily.

Polygon is an abstract class and has a vertex arrays m_ptlist which records vertexes orientated in counter-clockwise. FoldPolygon and CutPolygon is the inheritance of Polygon Class.

FoldPolygon has two data members, which are named m_pgChildArray and m_CPgArray. The m_pgChildArray stores the nodes as folding children of the polygon list. The m_CPgArray stores the results of clipping polygons generated by Cut-Polygon and the current FoldPolygon, the results are the instances of CutPolygon.

Class FoldPolygon and Class CutPolygon overwrites the Cut method and Merge method of Class Polygon. In the following, these two methods will be rewritten in accordance with different needs of instances.

3 Folding, Cutting and Unfolding Algorithms

Traditional creation of paper-cut involves three steps, which are folding paper, cutting out the paper and unfolding paper. We represent the multi-tree as the whole paper's interior data structure after the study of the process in Paper-cut. The Folding Algorithm, Cutting Algorithm and Unfolding Algorithm are all designed on the base of the multi-tree.

Paper's Folding Algorithm will be expounded in section 3.1. In section 3.2, two algorithms about Cutting Algorithm will be described. One is algorithm between Folding Polygon and Cutting Polygon, and the other one is between two Cutting Polygons. Last, the Reform Outline Algorithm will be described in the section 3.3.

3.1 Folding Algorithm

Folding Algorithm of the paper is implemented via Paper::UnFolding(). First, leaf nodes of the multi-tree must be gave out by travel. If the leaf node is folded, it will generate at least two new leaf nodes.

Our Folding Algorithm is illumed by Scan-line Algorithm [3] and it involves four steps, geometry center and intersection calculation, intersection list generation, grouping intersection into pairs and polygons generation, respectively.

The related assistant data structures are:

CrPntlist：Intersection list, which records the intersection points between polygon's edges and folding axis.

InsertP：Insert list, which contains the points in CrPntlist and points in m_ptlist, all points in Insert list must be oriented as the same as circuit of the m_ptlist.

Polygon::Folding(Axis* axis):

step I Calculating out the geometry center, as displayed by the yellow dot in Figure 3;

step II Finding out the intersections between the edges and axis, if none, exit, or else, add intersections in the CrPntlist;

step III As depicted in step 2 in Figure 3, polygon must be rotated until the point tagged zero is on the same side with the geometry center, and the point tagged one is in the different side with the geometry center;

step IV As depicted in step 3 in picture 3, generating InsertP, intersection points must be added in m_ptlist in the counterclockwise orientation.

step V Generating new polygons

a) Grouping intersection points into pairs. From i(i=0)th intersection point, doing grouping in turn. Finding the index j of the ith intersection point in InsertP, estimating the position of the point tagged as j+1 in InsertP to know if it is the same as the geometry center's position relative to folding axis. If the two points lies apart by the folding axis, then intersection points marked as i th and i+1 th must be grouped. Sequentially, still grouping intersection points form i+2 th point in Crpnlist. If they are not apart by axis, the current intersection point must be skipped, and i+1 th in CrPntlist must be considered. Repeat this step, until all the intersection points are grouped. This step is described in process of step 2 to step 3 in Figure 3;

b) Generating new polygons from all pairs of the grouped intersection points. From the i(i=0)th pair of grouped intersection points, the first intersection point, the points between two intersection points and the second intersection point are responsible for generation of polygon. They are sorted as contrary direction as their direction in InsertP and points between two intersections are marked as visited. After contrary direction sorting, corresponding points which relative to axis replace them and generate one new polygon P'. The new polygon P' is stored as the child of the current polygon in m_prChidArray. Doing this step unlit all pairs generate new polygons. This description is reflected in picture 3 step3 to setp4;

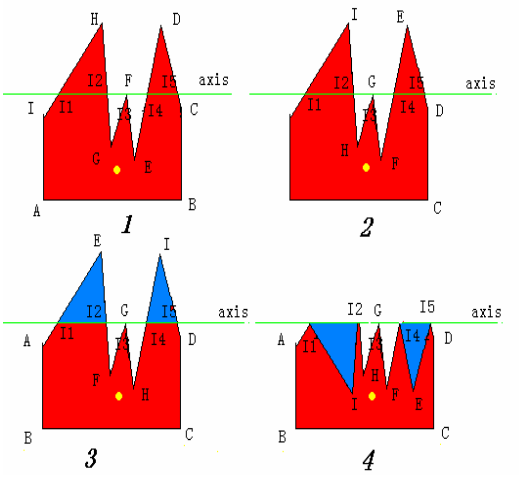


Fig. 3. The process of folding

c) Form the first point in InsertP, all point which are not visited and intersection points must be added as a circuit for generating new polygon P. New polygon is stored in m_prChidArray. This description is reflected in Figure 3 step3 to setp4;

As depicted in Figure 3, 1 to 2 is logic rotation;2 to 3 is generation of InsertP: A B C I5 D I4E F(I3) G I2 H I1 I and CrPntlist:I1 I2 I4 I5grouping into pairs;3 to 4 is new polygons generation.

The method Folding(Axis* axis) of FoldPolygon and CutPolygon inherit form Polygon::Folding(Axis* axis).Whenever there are new CutPolygon generated by fold, a specific relation between new FoldPolygon and new CuttPolygon are estimated by logic.

3.2 Cutting Algorithm upon Polygons

Algorithm based upon Polygons is consist by algorithm between two instances of CutPolygon, and instances of CutPolygon and FoldPolygon. The Reform Outline Algorithm shared some advantages of our Cutting Algorithm, but different in the some point, will be discussed in section 3.4.

Our algorithms in this section are the evolvement of classical Weiler-Atherton[4] (shortened form WA, as below), which deals with two kinds of polygon, one is called clip polygon and another one is called subject polygon that is be clipped, mainly on three steps:

step I Building two vertex lists of the two polygon, respectively;

step II Calculating intersection points of the two polygons. Intersection points list, two insert lists that are InsertS and InsertC, and one bidirectional list are be generated.

step III Clipping or Combination.

In the real situation of the Paper-cut, it is pretty obviously that one polygon's edge will overlap with another one's edge or vertex will lies on the edge of another. WA also gives out salvation aimed at this confused situation. Special vertexes and edges need two more consideration listed below:

1. Edges belonged to subject polygon and overlapped with clip polygon need not be taken into the calculation of intersection points;
2. Vertexes belong to edge of the subject polygon but lies on the edges of clip polygon, if the edge they belong to is in the polygon, then it must be taken as intersection points, or else, they are not.

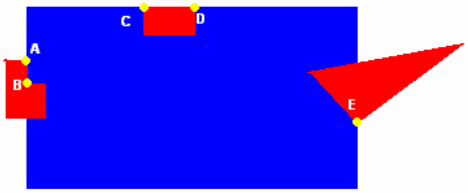


Fig. 4. Special situation and the yellow intersection points

As described in Figure 4, the subject polygons are colored by red, and the clip polygon is colored by blue. Taking point A as the intersection point, and not regarding point B as intersection point, Algorithm would manage two polygons correctly. Rectangle would be disposed by regarding points C and D as the intersection points. Also, triangle would be disposed correctly, as if the condition of regarding E as intersection point were taken.

In our Algorithm, clipping method of WA is applied between the instances of FoldPolygon and CutPolygon.

We considered instance of FoldPolygon as the subject polygon and instance of CutPolygon is clip polygon. Still, new polygons are generated via the WA are regard as the instances of CutPolygon. The new ones are stored in the m_CpgArray of the subject polygon.

Combination method of WA is applied between CutPolyongs. In some peculiar situation, after the combination, there would be some polygons appended "islands". Therefore, we add one step to handle this strange situation which never occurs in Paper-cut. We delete polygon oriented in clockwise and only reserve one polygon in counterclockwise. Figure 5 reflects what we describe above. Polygon i1c1c6i4s7s8s1s2 needs to be reserve and polygon i2s3s4s5i3c4c4 must be deleted.

3.3 Unfolding Algorithm

Fold, Cutting and Unfolding are the three steps of the Paper-cut. They work in sequence in practice. Accordingly, all the Algorithms are designed in this order. First, through the Folding Algorithm, it generates a multi-tree. Second, it is Cutting Algorithm. The leaf nodes of the tree and instances of CutPolygon are involved by this algorithm to generate some new polygons which regarded as instances of CutPolygon. Last, it is Unfolding Algorithm. In Unfolding Algorithm, some of the instances of

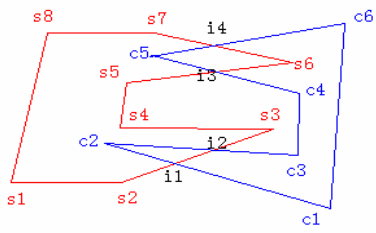


Fig. 5. Island

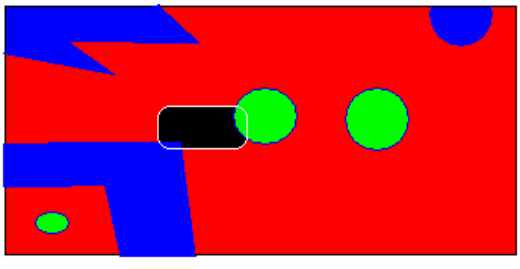


Fig. 6. Folding Paper and decorative pattern

CutPolygon need the affine transforms due to the fold and they are not suitable for Reform Outline Algorithm directly. It is because that some of them occupy the edges of initial polygon described paper's shape, and some does not. In another word, some of the polygons could reshape the paper's outline directly, while some of them are regarded as decorative patterns called genus of the initial shape in glossary of topology.

According Fig 6, the red rectangle is the initial outline of the paper. Blue polygons are the ones which occupied the initial outline's edge. Green ones and black ones are the genus according to initial polygon. The figuration of the final artistic work, in some ways, is the XOR between initial outline and Cutpolygon, actually. Some unwanted step could be skipped after the genus and none-genus are sorted.

Upon these reasons, we divide the results into two categories which are set A and set B. The set A contains polygon which are not genus according to initial outline. The genuses are aggregated in the set B. In Figure 6, the blue ones are in the set A, the rest of them belong to set B.

Unfolding Algorithm:

step I Producing the set A and set B;

step II Taking each element in B for combination with every element in A, if the combination is successful, redefine the elements both in A and B. Replacing the polygon inhere in A by the new one generated after the combination deleting the polygon inhere in B. Repeating this step until no more combination occurs, thus, two new sets A* and B* are gave out;

step III Do the Reform Outline algorithm until all elements in A* are taken in;

step IV Reserving the elements of set B appears in final outline.

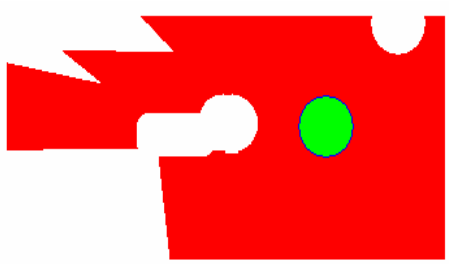


Fig. 7. Result

In Figure 6, after the setp II, the black polygon is absorbed in set A, then A* comes out. After setp IV, the smaller green genus at the left bottom is deleted, and the left one is reserved. Figure 7 is the final result.

It is Reform Outline Algorithm in setp III. This algorithm shares some inspirations with the WA. As we write below, taking the outline P as the subject polygon and all elements in A* as the clipping polygon：

step I Building two vertex lists of the two polygon respectively;

step II Calculating intersection points of the two polygons. Intersection points list, two insert lists that are InsertP and InsertC, and one bidirectional list are generated.

step III Reshape the outline:

a) Building a new temporary vertex list;

b) Selecting one vertex that has never been visited and is not the intersection point in InsertP, then add it into the temporary vertex list;

c) Adding the points in InsertP in the positive order (counter-clockwise), converting to InsertC when meet an intersection point, and then still add points but in negative order (clockwise); whenever meets an intersection points, current insert list must be converted. This step ended until the first point comes out again. Marking the current polygon generated by these points as the P';

d) Repeating the step a b c, until all the candidates suit step b are visited.

step IV For all P', convert the original P by P' which has the largest area.

If there is any special situation, such as vertexes or edges occupy other edge should follow the instructions gave out by WA.

In this section, we introduce our Unfolding Algorithm. Unfolding Algorithm is consisted by improved WA applied in combination between CutPolygon's instances and Reform Outline Algorithm. The initial outline was decided by the point in Paper::Init(Polygon* pg), alternatively, could be obtained through hand sketch or [5][6]. By [5][6], a consecutive points array will be obtained. We will refine consecutive points array by applying [8-12]. Some of the parametric curves are also refined by those algorithms. Vignettes are made by Bezier in manual or via automatic method.

4 Implementation and Discussion

In the environment of Windows operating system, using VS 2005 as the development platform we designed and developed the "two-dimensional paper folding and cutting system", the interface is showed in figure 6 below. According to many tests, it demonstrates that the algorithms we take are accurate and reliable, and it has many advantages:

a) 2-D processing, it's easy for user to operate;
b) The system is open, scalable and reusable.
c) The data supports SVG standard so that the output results can be further used for other image processing software to edit.
d) The system can deal with the primary material effect.

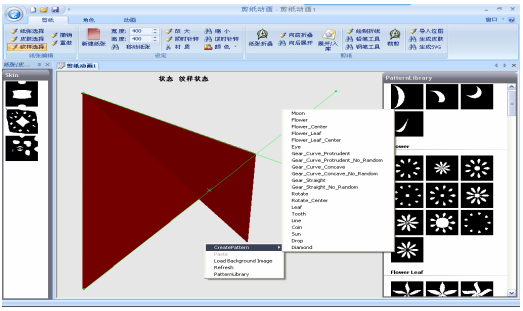


Fig. 8. On top of the interface is a control panel, In the middle is a work area, on the left side is the finished modeling, on the right side is the decorative patterns

Following is the works created by artists using our paper folding and cutting system.

It is an Olympic Mascot in Figure 10.First, as we introduced above, the artist use our system to make a semi finished articles, except the nose and eyes of the character, due to unwanted "island" in our algorithms. Secondly, with the second disposal in SVG Developer, the eyes and nose of this character are added as the pattern. All these efforts make it more vivid than before.

In the following researches, we will focus on the folding and cutting algorithms of paper with different materials. At present, our system can fold and cut papers of different colors. The system needs further improvement and expansions, so it can be polished to manage papers of more than one material effect and manage paper by more simulation algorithms. It means that we should take the needs of artists and ordinary users into consideration. From the following aspects we can try our best to achieve the perfect unity of art and computer simulation.

1. The realization of the paper coloring: including chromatic, coloring, color separation, color contrast, dislocation stack liner, overturned and spray color, and so on.
2. The realization of paper-cut and pasting. Those have not been achieved in our system.

Therefore, the realization of paper coloring, paper-cut and pasting will be the goals of my next step's work.

Fig. 9. Some artistic works

Fig. 10. Olympic Mascot

Acknowledgement

The author would like to acknowledge the staff of Digital Technology and Digital Art R&D Center and his supervisor prof. Minyong Shi. Also, this program is funded by Technology Department, State Administration of Radio Film and Television (Grant No. 2006-11).

References

1. Yu, J., Luo, G., Peng, Q.: Image Based Synthesis of Chinese Landscape Painting [J]. Journal of Computer Science and Technology 18(1), 22–28 (2003)
2. Rogers, D.F.: Procedure Elements for Computer Graphics. McGraw Hill, New York (1985)
3. Sutherland, I.E., Hodgman, G.W.: Reentrant polygon clipping. Comm. ACM 17(1), 28–352 (1974)
4. Weiler, K.J., Atherton, P.R.: Hidden surface removal using polygon area sorting. Computer Graphics 11(2), 214–222 (1977)
5. Donna, W., Mubarak, S.: A fast algorithm for active contours and curvature estimation. CV GIP: Image Understanding 55(1), 14–26 (1992)
6. Kass, M., Witkin, A., Terzopoulos, D.: Snake: active contour models. International Journal of Computer Vision 1(4), 321–331 (1998)
7. Riesenfeld, R.F.: An algorithm for high speed curve generation[A]. Computer Graphics and Image Processing[C] 3(4), 346–349 (1974)
8. Carstensen, C., Mühlbach, G., Schmidt, G.: De Casteljau's algorithm is an extrapolation method [J]. Computer Aided Geometric Design 12(4), 371–380 (1995)
9. de Boor, C.: Cutting corners always works [J]. Computer Aided Geometric Design 4(122), 125–131 (1987)
10. The, C.-H., Chin, R.T.: On the Detection of Dominant Pointson Digital Curves [J]. Trans on Pattern Analysis and Machine Intelligence 11(8) (1989)

A Sufficient Condition for Uniform Convergence of Stationary *p*-Subdivision Scheme

Yi-Kuan Zhang[1, 2], Ke Lu[3], Jiangshe Zhang[1], and Xiaopeng Zhang[2,*]

[1] School of Science, Xi'an Jiaotong University, Xi'an 710049, China
[2] LIAMA-NLPR (National Laboratory of Pattern Recognition), Institute of Automation, Chinese Academic of Science, Beijing 100080, China
[3] Graduate University of Chinese Academy of Science, Beijing, 100049 China
ykzhang@liama.ia.ac.cn, luk@gscas.ac.cn, jszhang@mail.xjtu.edu.cn, xpzhang@nlpr.ia.ac.cn

Abstract. Subdivision is a convenient tool to construct objective curves and surfaces directly from given scattered points. Stationary *p*-subdivision schemes are highly efficient in the acquisitions of curve/surface points in shape modeling. The features of supported set of nonnegative mask of uniform convergent stationary subdivision schemes are important to their theoretic researches and applications. According to the properties of supported set of the nonnegative mask, a sufficient condition for uniform convergence of stationary *p*-subdivision scheme is presented. This condition is proved with two propositions and spline function. The contribution of this work is that the convergence of a stationary *p*-subdivision scheme can be judged directly. This direct judge is in favor of applications of this scheme.

Keywords: geometric modeling, stationary *p*-subdivision, uniform convergence, contractility, spline function.

1 Introduction

Stationary subdivision schemes arise from modeling and interrogation of curves and surfaces, image decomposition and reconstruction, and the problems of constructing compact supported wavelet basis etc. [1, 9]. These schemes are being developed in geometric modeling with great potentiality in CAD/CAM, computer graphics, image processing, etc. [1-11, 14, 15]. Stationary subdivision schemes are widely used in mechanical CAD, garment CAD, jewellery CAD, and applied in computer graphics. They also play important roles in image coding, signal processing, and the construction of basis function of compact supported orthogonal wavelets by using multiresolution analysis [1-3, 17-23]. They are also important in fractal and its generation by computer in particularly [1, 2, 4, 12]. Stationary subdivision schemes are used to construct the required curves and surfaces from scattered data directly through stated subdivision rules. Moreover the theoretical contribution of this approach consists in their tight combination of three research disciplines: spline functions, wavelets and

* Corresponding author.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 719–727, 2008.

fractals [1-4, 9, 10, 12]. Therefore, the research of stationary subdivision schemes, especially its convergence, is significant in theoretical research and shape modeling [2, 3, 5-13]. The idea and approaches of stationary subdivision schemes are still effective in subdivision surfaces [24, 25,26] and the constructions of compactly supported orthogonal wavelets basis and fractal [11, 12, 13, 16, 24].

The systematic development of the basic mathematical principles and concepts associated with stationary 2-subdivision schemes is presented in [1]. The structure of these algorithms in a multidimensional setting and convergence issue are researched systematically. The complete theoretical system is constructed. The analytic structure of limit curves and surfaces generated by these algorithms is revealed [1, 16].

The extension of stationary 2-subdivision to stationary p-subdivision scheme is presented in [9]. Some properties of convergence of such schemes are described through Fourier analysis, functional analysis and spline function. A sufficient condition of the uniform convergence of the stationary p-subdivision scheme is discovered in [10] through a special polygon, δ-control polygon.

The problem is important of how to use this kind of subdivision schemes to generate curves and surfaces in computer graphics [1-8, 18-21]. The convergence of stationary subdivision schemes is a key problem in the theory of stationary subdivision scheme and their applications [1, 9, 10]. Finding the features of the support set of nonnegative mask $\mathbf{a} = \{a_\alpha : \alpha \in \mathbf{Z}^s\}$ [1, 9, 10] has an important value in theoretical researches and practical applications, because the convergence of these algorithms can be judged directly in the construction of curves and surface. So, the sufficient conditions of the uniform convergence of stationary p-subdivision schemes based on the supported set of mask may promote theoretical researches and practical applications [1, 9, 10].

A sufficient condition of the uniform convergence of stationary p-subdivision schemes is presented in this paper using contractility and spline function. This work is based on three aspects: the nonnegative mask and its support set of stationary p-subdivision schemes, some definitions and properties of stationary p-subdivision schemes presented in [9, 10], and the works of [1-5].

Here is the main theorem of this paper.

Theorem. The stationary p-subdivision scheme

$$\lambda^0 = \lambda, \quad \lambda^m = S\lambda^{m-1}, \qquad m = 1, 2, \cdots$$

defined in (2) is uniformly convergent, if the positive mask $\mathbf{a} = \{a_\alpha : \alpha \in \mathbf{Z}^s\}$ supported on Ω satisfies

$$\sum_{\beta \in \mathbf{Z}^s} a_{\alpha - p\beta} = 1, \quad \alpha \in \mathbf{Z}^s,$$

where

$$\Omega := \mathbf{Z}(A) \bigcap \mathbf{Z}^s \text{ , zonotope } \mathbf{Z}(A) := \{Au : u \in \prod_{i=1}^{s} [l_i, u_i]\}, \ l_i + p - 1 < u_i \text{ , } i = 1, \cdots, s \text{ .}$$

2 Preliminaries and Propositions

Six definitions and two propositions are introduced in order to prove of above theorem.

Definition 1. Let s be a fixed natural number and $\mathbf{Z}^s$ the integer lattice, and $\mathbf{a}=\{a_\alpha : \alpha \in \mathbf{Z}^s\}$ be the fixed real scalar sequences having finitely supported set supp$\mathbf{a}=\{\alpha : a_\alpha \neq 0\}$. A stationary p-subdivision operator S is defined as

$$S : \ell^\infty(\mathbf{Z}^s) \to \ell^\infty(\mathbf{Z}^s) \tag{1}$$

by

$$(S\lambda)_\alpha = \sum_{\beta \in \mathbf{Z}^s} a_{\alpha - p\beta}\lambda_\beta , \quad \lambda \in \ell^\infty(\mathbf{Z}^s),$$

where $p>1$ is a fixed natural number, and λ is point sequence.

Definition 2. Let S be any stationary p-subdivision operator defined in (1), the following iteration scheme

$$\lambda^0 = \lambda, \quad \lambda^m = S\lambda^{m-1}, \qquad m = 1, 2, \cdots \tag{2}$$

is defined as a stationary p-subdivision scheme. $\mathbf{a}=\{a_\alpha : \alpha \in \mathbf{Z}^s\}$ is referred to as the mask of the stationary p-subdivision scheme S. λ is called as the control polygon of S. In fact, λ_α is a vertex of the control polygon λ.

Definition 3. The stationary p-subdivision scheme (2) is said to be convergent for $\lambda \in \ell^\infty(\mathbf{Z}^s)$ if there exists a continuous function $f_\lambda \in \mathbf{C}^0(\mathbf{R}^s)$, such that

$$\lim_{m\to\infty} \left\| f_\lambda(\frac{\bullet}{p^m}) - \lambda^m \right\|_\infty = 0 . \tag{3}$$

Definition 4. The p-subdivision scheme (2) is said to be uniformly convergent if there exists a continuous function $f_\lambda \in \mathbf{C}^0(\mathbf{R}^s)$ for all $\lambda \in \ell^\infty(\mathbf{Z}^s)$, such that

$$\lim_{m\to\infty} \left\| f_\lambda(\frac{\bullet}{p^m}) - \lambda^m \right\|_\infty = 0 . \tag{4}$$

Stationary p-subdivision algorithms (1) actually have p^s different subdivision rules, and the norm is defined as $\|\lambda\|_\infty = \sup_{\alpha \in \mathbf{Z}^s} |\lambda_\alpha|$ in $\ell^\infty(\mathbf{Z}^s)$ in the above definitions. The control polygon λ is represented as scalar-valued, i.e. $\lambda \in \ell^\infty(\mathbf{Z}^s)$, in this paper, since the influence on λ of stationary p-subdivision scheme S in (1) and (2) is performed as that of coordinate components of vertices.

The basic difference of stationary p-subdivision schemes and stationary 2-subdivision schemes is that stationary p-subdivision schemes have p^s different rules while stationary 2-subdivision schemes have 2^s different rules. If p>2 and the two kind of stationary subdivision schemes in (2) are convergent, stationary p-subdivision schemes can be used to generate curves or surfaces by with fewer iterative steps than stationary 2-subdivision schemes, so stationary p- subdivision schemes are subdivision schemes having a faster convergence speed and a higher efficiency in curve/surface modeling.

Stationary p-subdivision schemes and their some basic convergent properties are presented in [9]. A sufficient condition for uniform convergence of stationary p-subdivision schemes is given in [10] by using a special control polygon of δ-control polygon.

$D:\ell^{\infty}(\mathbf{Z}^s)\to\mathbf{R}^{+}\bigcup\{0\}$ is thought as a no-trivial non-negative functional in the following description.

Definition 5. A stationary p-subdivision operator S is said to be contractive relative to D if there exists a constant number γ ($0<\gamma<1$) for the subdivision operator S defined by (1), such that

$$D(S\lambda)\le\gamma D(\lambda),\quad \lambda\in\ell^{\infty}(\mathbf{Z}^s). \tag{5}$$

Suppose $\mu\in\mathbf{R}^s$ is a fixed vector not necessarily a lattice point, and $\sup p\mathbf{a}\subseteq\Omega:=(\mu+\mathbf{\Gamma})\bigcap\mathbf{Z}^s$, where $\mathbf{\Gamma}\subseteq\mathbf{R}^s$ be a balanced convex closed set which corresponding Minkowski functional is ρ, then $y\in\mathbf{\Gamma}$ if and only if $\rho(y)\le 1$.

Definition 6. For any control polygon $\lambda\in\ell^{\infty}(\mathbf{Z}^s)$,

$$D_{\rho}(\lambda):=\sup_{\substack{\rho(\alpha-\beta)<2\\ \alpha,\beta\in\mathbf{Z}^s}}\left|\lambda_{\alpha}-\lambda_{\beta}\right| \tag{6}$$

is defined as a diameter of λ.

The convergent condition of stationary p-subdivision scheme will be discussed in the following under the condition of that the support of mask $\mathbf{a}=\{a_{\alpha}:\alpha\in\mathbf{Z}^s\}$ is the union of $\mathbf{Z}^s$ and a special zonotope $\mathbf{Z}(A):=\{Au:u\in\prod_{i=1}^{s}[l_i,u_i]\}$, where A is a $s\times s$ integer matrix, and $\det A=-1$. The following two propositions are used to prove the main theorem.

Proposition 1. Let $\mathbf{a}=\{a_{\alpha}:\alpha\in\mathbf{Z}^s\}$ be any mask satisfying following conditions:

$$a_{\alpha}\ne 0,\ implies\quad \alpha\in\Omega, \tag{7}$$

$$\sum_{\beta\in\mathbf{Z}^s}a_{\alpha-p\beta}=1,\quad for\ \forall\alpha\in\mathbf{Z}^s; \tag{8}$$

and $D_{\rho}(\lambda)$ be functional defined by (6) on $\ell^{\infty}(\mathbf{Z}^s)$. Then stationary p-subdivision operator S defined by (1) satisfies:

$$D(S\lambda)\le\gamma_{\rho}D_{\rho}(\lambda)\ , \tag{9}$$

Where

$$\gamma_{\rho}=\frac{1}{2}\max_{\rho(\sigma-\delta)<2}\sum_{\beta\in\mathbf{Z}^s}\left|a_{\sigma-p\beta}-a_{\delta-p\beta}\right|. \tag{10}$$

Lack of space forbids the proof of this proposition here.

Proposition 2. Under the following three conditions

(i) B be a p-subdivision operator which has finitely supported mask $\mathbf{b} = \{b_\alpha : \alpha \in \mathbf{Z}^s\}$ and stable refinable function ψ, and the corresponding stationary p-subdivision scheme is uniform convergent: $(B\lambda)_\alpha := \sum_{\beta \in \mathbf{Z}^s} b_{\alpha - p\beta}\lambda_\beta$. Where ψ is a stable refinable function means that for refinable function ψ there exists a positive constant $C_1 > 0$ such that

$$C_1 \|\lambda\|_\infty \le \left\| \sum_{\alpha \in \mathbf{Z}^s} \lambda_\alpha \psi(\bullet - \alpha) \right\|_\infty \le C_{21} \|\lambda\|_\infty , \tag{11}$$

Where $C_2 := \left\| \sum_{\alpha \in \mathbf{Z}^s} |\psi(\bullet - \alpha)| \right\|_\infty$.

(ii) Stationary p-subdivision operator S defined by (1) is contractive relative to functional D.

(iii) There exists a constant C, such that

$$\|S\lambda - B\lambda\|_\infty \le C \cdot D(\lambda), \quad \lambda \in \ell^\infty(\mathbf{Z}^s) . \tag{12}$$

The following two conclusions can be obtained

(a) The stationary p-subdivision scheme determined by S is uniformly convergent.

(b) If the condition (11) is replaced by the following condition that there exist a constant C such that

$$|\lambda_\alpha - \lambda_\beta| \le C \cdot D(\lambda), \quad \beta - \alpha \in (\sup p\psi)^\circ , \tag{13}$$

the stationary p-subdivision scheme determined by S is also uniformly convergent.

Proof: For the definitions in (1) and (2), we can obtain that [10]

$$\lambda_\alpha^m = (S^m \lambda)_\alpha = \sum_{\beta \in \mathbf{Z}^s} a^m_{\alpha - p^m \beta}\lambda_\beta \text{ and } a_\alpha^m = (Sa^{m-1})_\alpha = \sum_{\beta \in \mathbf{Z}^s} a_\beta^{m-1} a_{\alpha - p\beta} . \tag{14}$$

If define f_λ^m as follows

$$f_\lambda^m(x) = \sum_{\alpha \in \mathbf{Z}^s} \lambda_\alpha^m \psi(p^m x - \alpha), m = 1, 2, \cdots ,$$

the conclusion of proposition 2 can be proved with following inequality [10]

$$\left\| f_\lambda(\frac{\bullet}{p^m}) - \lambda^m \right\| \le \left\| f_\lambda(\frac{\bullet}{p^m}) - f_\lambda^m(\bullet) \right\| + \left\| f_\lambda^m(\bullet) - \lambda^m \right\| .$$

The detailed proof of this proposition is omitted here.

3 The Proof of the Main Theorem

Now we give the proof of theorem presented in this paper on the base of the proposition 1 and proposition 2. In the proof, above definition of contractility and spline function will be used.

Proof: (i) Let $\Omega := \mathbf{Z}(A) \bigcap \mathbf{Z}^s$, then it is a hyperrectangle according to the definition of $\mathbf{Z}(A)$. So we can suppose that $\mathbf{Z}(A) := \{Au : u \in \prod_{i=1}^{s} [l_i^1, u_i^1]\}$. Then according to the hypothesis of that Ω is the support of mask $\mathbf{a}$, we know that:

$$a_{\sigma - p\beta}, a_{\delta - p\beta} > 0 \quad \Leftrightarrow \quad l_i^1 \le \sigma_i - p\beta_i, \delta_i - p\beta_i \le u_i^1, i = 1, 2, \cdots, s \tag{15}$$

So, if let

$$\mu = \frac{1}{2}(u^1 - l^1),\ u^1 = (u_1^1, u_2^1, \cdots, u_s^1),\ l^1 = (l_1^1, l_2^1, \cdots, l_s^1),$$

$$\rho(x) := 2 \max_{1 \le i \le s} \left| \frac{\sigma_i - \delta_i}{u_i^1 - l_i^1} \right|,\ x = (x_1, x_2, \cdots, x_s),$$

and $\mathbf{\Gamma}$ is the set determined by Minkowski functional $\rho(x)$, as a result , we know that $\Omega = (\mu + \mathbf{\Gamma}) \bigcap \mathbf{Z}^s$ from the definition of μ and $\mathbf{\Gamma}$, so $a_\alpha > 0$ if and only if $\alpha \in \Omega = (\mu + \mathbf{\Gamma}) \bigcap \mathbf{Z}^s$.

Moreover, the mask $\mathbf{a} = \{a_\alpha : \alpha \in \mathbf{Z}^s\}$ satisfies (8) because of known conditions. So, for $D_\rho(S\lambda)$ defined with (6), we conclude that $D_\rho(S\lambda) \le \gamma_\rho D(\lambda)$ according to proposition 1. Thus from that a_α $(\alpha \in \mathbf{Z}^s)$ is positive on Ω, it follows that

$$\gamma_\rho = \frac{1}{2} \max_{\rho(\sigma - \delta) < 2} \sum_{\beta \in \mathbf{Z}^s} \left| a_{\sigma - p\beta} - a_{\delta - p\beta} \right| \le \frac{1}{2} \max_{\rho(\sigma - \delta) < 2} \left(\sum_{\beta \in \mathbf{Z}^s} a_{\sigma - p\beta} + \sum_{\beta \in \mathbf{Z}^s} a_{\delta - p\beta} \right) = 1$$

Therefore, if find a β satisfying (15) whenever $\left|\sigma_1 - \delta_1\right| < u_i^1 - l_i^1 = u_i - l_i$. Then for such β $\left|a_{\sigma - p\beta} - a_{\delta - p\beta}\right| < a_{\sigma - p\beta} + a_{\delta - p\beta}$ is true. Thus $\gamma_\rho < 1$.

(ii) To determine the β satisfying above requirement in the following.

From the inequalities in (15) above, it may be known that in order to make the $a_{\sigma - p\beta} > 0,\ a_{\delta - p\beta} > 0$ true the subscripts $\sigma - p\beta,\ \delta - p\beta$ should satisfy:

$$l^1 \le \sigma - p\beta \le u^1, l^1 \le \delta - p\beta \le u^1. \tag{16}$$

So the expression $\left|\sigma - \delta\right| < \left|u^1 - l^1\right|$ is true since $\left|\sigma_i - \delta_i\right| < \left|u_i^1 - l_i^1\right|,\ i = 1, 2, \cdots, s$. So

$$\rho(\sigma - \delta) := 2 \max_{1 \le i \le s} \left| \frac{\sigma_i - \delta_i}{u_i^1 - l_i^1} \right| < 2.$$

Without loss of generality, let $\sigma > \delta$, then $0 \le \sigma - \delta < u^1 - l^1$, and so $\sigma - u^1 < \delta - l^1$. Now to solve the $p\beta$ from the expression (16), the result $\sigma - u^1 \le p\beta \le \delta - l^1$ is obtained. Therefore there always exists a integer in interval $[\sigma - u^1, \delta - l^1]$, so that this integer is $p\beta$, so the β can be find out according to the each coordinate component, and such $\beta \in \mathbf{Z}^s$ is satisfies all the inequalities (15). Therefore the operator S has following property known from the conclusion (i):

$$D(S\lambda) \le \gamma_\rho D_\rho(\lambda) \text{ , and } 0 < \gamma_\rho < 1.$$

(iii) To construct an operator B defined in (1), which has finite support $\mathbf{b}=\{b_\alpha : \alpha\in \mathbf{Z}^s\}$ and refinable function ψ, and the B makes the corresponding stationary p-subdivision scheme be uniform convergent and satisfying (13).

Firstly, let $\varphi_1(t)$ is a B-spline function of degree one: $\varphi_1(t)=\begin{cases}1-|t| & ,|t|\le 1\\ 0 & ,others\end{cases}$. Now let $\varphi(\mathrm{x})=\prod_{i=1}^{s}\varphi_1(x_i)$, $x=(x_1,x_2,\cdots,x_s)$, if let $\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_s)$, then

$$\sum_{\alpha\in\mathbf{Z}^s}\varphi(x-\alpha)=\sum_{\alpha_1,\alpha_2,\cdots,\alpha_s\in\mathbf{Z}^s}\varphi_1(x_1-\alpha_1)\varphi_1(x_2-\alpha_2)\cdots\varphi_1(x_s-\alpha_s)=\prod_{i=1}^{s}\sum_{\alpha\in\mathbf{Z}^s}\varphi_1(x_i-\alpha_i)$$

Considering $\sum_{\alpha\in\mathbf{Z}^s}\varphi_1(x_i-\alpha_i)=1$, so $\sum_{\alpha\in\mathbf{Z}^s}\varphi(x-\alpha)=1 \quad x\in\mathbf{R}^s$ (17)

Moreover if we let

$$b_\alpha=b_{\alpha_1}b_{\alpha_2}\cdots b_{\alpha_s},\ \alpha=(\alpha_1,\alpha_2,\cdots,\alpha_s),\ b_j=\begin{cases}1-1/p|j| & ,\ j\in\{-1,0,1\}\\ 0 & ,\ others\end{cases},$$

then the φ satisfies the p-scale equation: $\varphi(x)=\sum_{\alpha\in\mathbf{Z}^s}b_\alpha\varphi(px-\alpha)\quad ,x\in\mathbf{R}^s$. (18)

Now select $\eta=(\eta_1,\eta_2,\cdots,\eta_s)$, such that $l_i<\eta_i<u_i \quad ,i=1,2,\cdots,s$, and let $\psi(x)=\varphi(x-\eta)$, then $\sup p\mathbf{b}\subseteq\sup p\mathbf{a}$, and mask $\mathbf{b}$ is associated to ψ.

Since $\sum_{\beta\in\mathbf{Z}^s}a_{\alpha-p\beta}=1=\sum_{\beta\in\mathbf{Z}^s}b_{\alpha-p\beta},\ for\ \alpha\in\mathbf{Z}^s$, then there is following conclusion:

$$(S\lambda)_\alpha-(B\lambda)_\alpha=(\sum_{\beta\in\mathbf{Z}^s}a_{\alpha-p\beta}\lambda_\beta-\sum_{\beta\in\mathbf{Z}^s}Ca_{\alpha-p\beta})+(\sum_{\beta\in\mathbf{Z}^s}Cb_{\alpha-p\beta}-\sum_{\beta\in\mathbf{Z}^s}b_{\alpha-p\beta}\lambda_\beta),$$

where C is an arbitrary constant, B is determined by mask $\mathbf{b}=\{b_\alpha : \alpha\in\mathbf{Z}^s\}$.

For $\forall\beta\in\mathbf{Z}^s$, we can chosen proper C, to make $|\lambda_\beta-C|\le \frac{1}{2}D(\lambda)$ true. So,

$$|(S\lambda)_\alpha-(B\lambda)_\alpha|=\left|\sum_{\beta\in\mathbf{Z}^s}(\lambda_\beta-C)(a_{\alpha-p\beta}-b_{\alpha-p\beta})\right|\le D(\lambda)\quad ,\quad \alpha\in\mathbf{Z}^s. \quad (19)$$

The hypothesis conditions in proposition 2 are all true as shown in expressions (17), (18), and (19). Thus the theorem is true.

Acknowledgments. This work is supported in part by National Natural Science Foundation of China under Grant No.60672148, 60602062, and in part by Beijing Municipal Natural Science Foundation under Grant No. 4062033.

References

1. Cavareta, A.S., Dahmen, W., Micchelli, C.A.: Stationary subdivision. In: Memoirs of American Mathematical Society, vol. 93(453), pp. 1–185. American Mathematical Society, Providence (1991)
2. Prautzch, H., Micchelli, C.A.: Computing curves invariant under halving. Computer Aided Geometric Design 4, 140–321 (1987)

3. Micchelli, C.A., Prautzch, H.: Computing surfaces invariant under subdivision. Computer Aided Geometric Design 4, 321–328 (1987)
4. Micchelli, C.A., Prautzch, H.: Uniform refinement of curves. Linear Algebra and Its Applications 114/115, 841–870 (1989)
5. Chaikin, G.M.: An algorithm for high speed curve generation. Computer Graphics and Image Processing 3, 346–349 (1974)
6. Riesenfeid, R.E.: On Chaikin algorithm. Computer Graphics and Image Processing 4, 304–310 (1975)
7. Zhijie, C.: Convergence error estimation and some properties off four-point interpolation subdivision scheme. Computer Aided Geometric Design 12(15), 459–468 (1995)
8. Dyn, N., Gregory, J.A., Levin, D.: A 4-point interpolatory subdivision scheme for curve design. Computer Aided Geometric Design 4, 257–268 (1987)
9. Zhang, Y., Zhang, X., Tian, J.: Stationary p-Subdivision Algorithm (in Chinese). Journal of Engineering Graphics 22(3), 92–101 (2001)
10. Zhang, Y., Gong, S., Dou, J.: A Study on the Convergence of p-Subdivision Schemes (in Chinese). Computer Applications and Software 19(9), 43–45 (2002)
11. Dyn, N., Gregory, J.A., Levin, D.: Analysis of uniform binary subdivision schemes for curve design. Constructive Approximation 7(2), 127–147 (1991)
12. Schaefer, S., Levin, D., Goldman, R.: Subdivision Schemes and Attractors. In: Desbrun, M., Pottmann, H. (eds.) Eurographics Symposium on Geometry Processing 2005. Eurographics Association 2005, Aire-la-Ville, Switzerland. ACM International Conference Proceeding Series, vol. 255, pp. 171–180 (2005)
13. Song, L., Xinlong, Z.: Multivariate refinement equation with nonnegative masks. Science in China Series A: Mathematics 49(4), 439–450 (2006)
14. Hongqing, Z., Jun, H., Zhenglin, Y., Guohua, P., Shuili, R.: 4-point Subdivision Scheme with Three Parameters. Progress in Intelligence Computation & Applications. In: The 1st International Symposium on Intelligence Computation and Applications, ISICA 2005, vol. 1, pp. 409–412. Springer, Heidelberg (2005)
15. Micchelli, C.A., Sauer, T.: Sobolev Norm Convergence of Stationary Subdivision Schemes. In: Le Méhauté, A., Rabut, C., Schumaker, L.L. (eds.) Surfaces Fitting and Multiresolution Methods, pp. 245–260. Vanderbilt University Press, Nashville (1997)
16. Xinlong, Z.: Characterization of Convergent Subdivision Schemes. Analysis in Theory and Applications 14(3), 11–24 (1998)
17. Liu, W., Kondo, K., Zhang, L.: An Approximating Subdivision Scheme by Cutting and Upheaving the corner of Control Polygon. In: Proceedings of the 11th International Conference on Geometry and Graphics, pp. 134–138. Moscow State University of Technology Press, Moscow (2004)
18. Farin, G.: Curves and surfaces for computer aided geometric design: A Practical Guide, 2nd edn. Academic Press, Inc., Boston (1990)
19. Denis, Z., Schroder, P. (eds.): Subdivision for modeling and animation. SIGGRAPH 2000 Course #23 Notes. ACM Press, New York (2000)
20. Peter, S., Denis, Z.: Subdivision for modeling and animation. SIGGRAPH 1998 Course #36 Notes. ACM Press, New York (1998)
21. Breen, D.E.: IEEE Computer Graphics and Applications. In: Computer Graphics in Textiles and Apparel, vol. 16(5), pp. 26–27. IEEE Computer Society, Washington, DC (1996)
22. Peter, S., Sweldens, W. (eds.): Wavelets in computer graphics. SIGGRAPH 1996, Course #13 Notes. ACM Press, New York (1996)

23. Felmi, C., lortiz, M., Schröder, P.: Subdivision surfaces: A new paradigm for thin-shell finite element analysis. International Journal for Numerical Methods in Engineering 47(12), 2039–2072 (2000)
24. Dyn, N., Gregory, J.A., Levin, D.: Analysis of uniform binary subdivision schemes for curve design. Constructive Approximation 7(1), 127–147 (1991)
25. De Villiers, J.: On refinable functions and subdivision with positive masks. Advances in Computational Mathematics 24(1-4), 281–295 (2006)
26. Chen, F., Ding, Y., Liu, J., Wei, D.: A Novel Non-stationary Subdivision Scheme for Geometric Modeling. In: Proceedings of the Fourth International Conference on Computer and Information Technology (CIT 2004), vol. 00, pp. 748–752. IEEE Computer Society, Washington (2004)

Model and Animate Plant Leaf Wilting

Shenglian Lu[1,2], Xinyu Guo[2], Chunjiang Zhao[2,*], and Chengfeng Li[2]

[1] Laboratory of Digital Agriculture, Shanghai Jiaotong University, Shanghai 200040, China
[2] National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China
{lusl, guoxy, zhaocj, licf}@nercita.org.cn

Abstract. We describe a venation skeleton-driven method for modeling and animating plant wilting. The proposed approach uses a representation of a three-dimensional skeleton for a leaf blade. Firstly, the leaf skeleton is used to generate a detail mesh for the leaf surface, and a venation skeleton is also generated interactively from the leaf skeleton. Each vein in the venation skeleton consists of a segmented vertices string. Secondly each vertex in the leaf mesh is banded to the nearest vertex in the venation skeleton. We then deform the venation skeleton by controlling the movement of each vertex in the venation skeleton by rotating it around a fixed vector. Finally the leaf mesh is mapped to the deformed venation skeleton, as such the deformation of the mesh follow the deformation of the venation skeleton. We apply our techniques to simulate wilting plant leaves resulted from biological response.

Keywords: plant leaf, skeleton-based shape deformation, motion simulation, natural phenomena simulation.

1 Introduction

The literature of realistic modeling of leaves has a long history in computer graphics. This is, in part, due to their either beautiful or colorful images. Partly they have a strong visual effect on the audience. Many techniques have been proposed for modeling the geometry or shape of leaves. Most of these methods, however, just describe the shape of leaves in regular way.

Some researchers have endeavored to generate curled shapes of plant leaves. Prusinkiewicz *et al.* [1] provided a detailed representation of combining interaction and parameterized algorithms for realistic plant modeling and scene creation involving plants, including curled leaves. Mündermann *et al.* [2] proposed a method for modeling lobed leaves. Effects of curled leaves could be generated by using free-deformation in their framework. Recently, Sung *et al.* [3] proposed an interactive method for modeling curled leaf surface. But it could involve excessive manual interactions for generating a desired curled shape of leaf by using their method. Studies on curvature of plant leaves from biophysical viewpoints have raised the question of what role, if any, genes play in the control of curvature [4]. Yet others study wave or wrinkle pattern in leaves with physical analysis [5]. But these may beyond our focus in this paper.

* Corresponding author.

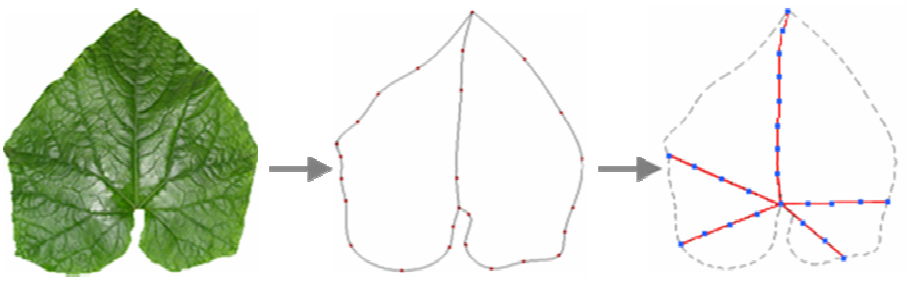
Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 728–735, 2008.

Additionally, there has been a great deal of previous work on the simulation motions of plant, including plant growth, motion in the wind and so on, such as the work demonstrated in [6]. Some physically-based models had been used to create natural plant shape [7, 8]. Recently Wang *et al.* [9] had simulated physically the growth of a plant leaf, the physical model used in the simulation is the governing equations of fluid mechanics – the Navier-Stokes equations. But the leaf model in their simulator was 2D.

Basing on the fact that less work has focused on modeling leaf surface deformation and simulating subtle behaviors of plant, such as wilting of leaves suffering from insufficient water supply, this paper presents a venation skeleton-based deformation method for plant leaves, and aims to develop an approximately kinematic model of leaf for simulating motions of plant leaves, especially wilting. The leaf skeleton plays two roles. It is used to generate a venation skeleton for later deformation. A geometric mesh for the leaf surface is also constructed from the leaf skeleton, and each vertex in the mesh is mapped to its nearest vertex in the venation skeleton, as such the geometric mesh is deformed according to the deformed venation skeleton. Applications of our approach to simulate wilting plant leaves with realistic results illustrate that the proposed model is flexible and effective.

2 Generating Venation Skeleton

The venation structure plays a major biological role in determining the leaf surface shape and controlling its deformation, as such we use it to control the deformation of a leaf blade. To generate the venation skeleton, we consider currently an interactive method but seem to be automated. Fig. 1 illustrates the process of generating venation skeleton of a leaf.



(a) scanned image (b) leaf skeleton (c) venation skeleton

Fig. 1. Process of generating venation skeleton

We use a representation of a leaf skeleton consisting of two boundary curves and a mid-vein curve which consist of feature points, as shows in Fig.1 (b). These boundary curves can be reconstructed from feature points in the boundary of a leaf, while these feature points can be extracted automatically from a scanned digital image by using a standard edge detection algorithm, or obtained by using 3D digitizer.

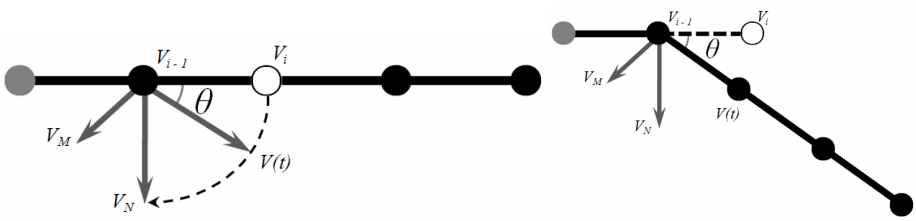
Generating skeleton of an object is a complex problem in computer graphics. Practical extraction of the skeleton of a 3D shape is usually based on 3D Voronoi diagram

techniques [10]. For our needs, we develop an interface for generating venation skeleton from a leaf skeleton interactively. As Fig.1 shows, the leaf skeleton can be obtained from a scanned image, a venation skeleton is then generated from the leaf skeleton. Each vein is segmented. Fig.1 (c) (the red line) shows a generated venation skeleton consisting of one mid-vein and four secondary veins, the blue vertices string segment each vein into several line sections. The process of generating venation skeleton involve several interactive manipulation including defining the start point and end point for each vein, and setting parameter for segmenting each vein. And users can decide how the mid-vein crosses the leaf skeleton, and how many secondary veins are attached to the mid-vein.

3 Venation Skeleton-Driven Leaf Surface Deformation

3.1 Mechanism for Controlling Movement of the Venation Skeleton

Skeleton-based methods have been widely used in computer animation and computer modeling for mesh deformation. Traditional skeleton-based methods commonly required a tedious process of weight selection to obtain satisfactory results. Note that the natural deformation of plant leaves is different from the deformation of humans' or animals' organs, we can constraint the movement of vertices in the venation skeleton. The most regular motions of plant leaves which can be seen by naked eyes are curling and wilting. The major goal of our approach is to develop an approximately kinematic model of plant leaf for simulating these motions.



(a) The movement locus of a skeletal joint (b) Child joints follow the parent's movement

Fig. 2. Demonstrates how the venation skeleton works

For our needs, we restrict the movement of each vertex in the venation skeleton by rotating it around a fixed vector. Fig.2 illustrates how a venation skeleton works. For convenience, the venation skeleton includes a vein, which consists of four segments. The light black vertex serves as the root node. Note that the movement of a leaf blade is always downward in the course of wilting, as such each joint in the skeleton segment does spherical movement. Take vertex V_i for example, it will align gradually vector $V_{i-1}V_N$ in the course of wilting, in which $V_{i-1}V_N$ is a downward vector reverse to Z axis, while vector $V_{i-1}V_M$ is perpendicular to the plane contains vector $V_{i-1}V_i$ and $V_{i-1}V_N$. The movement of vertex V_i can be looked as V_i rotating round vector $V_{i-1}V_M$.

To obtain a motion sequence of a vertex in a skeleton, a simple method is to rotate the vertex with a fixed angle, such as θ in Fig. 2(a), and the angle is commonly given

by the users. We have mentioned before that a vertex in the skeleton will do spherical rotation during the wilting of its corresponding leaf surface. As Fig. 2(a) illustrates, we can calculate the new position of vertex V_i by the following parametric equation:

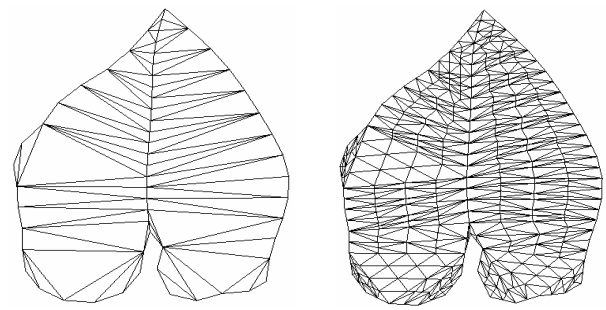
$$V(t) = k(t) * (V_i + t * (V_N - V_i)) \quad (1)$$

Where $k(t) = |V_i| / |V_i + t * (V_N - V_i)|$ $(0 \leq t \leq 1)$. Further, the motion sequence of V_i can be obtained by increasing parameter t. This may simplify the rotation operation.

As we know, all child segment of a vertex in the skeleton will follow the movement of the vertex, this can be done by passing a displacement and rotation angle to its child vertices when the vertex is rotated. Fig. 2(b) gives results after moving V_i to $V(t)$.

3.2 Constructing Leaf Surface

We have constructed the leaf skeleton with two boundary curves and a mid-vein curve as shows in Fig.1 (b). To mesh the void within these curves, we employ Delaunay triangulation scheme. The reason for using Delaunay triangulation is to deal with the problem of existing concave region in the leaf skeleton. When using Delaunay triangulation, we can use directly the feature points in the mid-vein and silhouette, or extract a series of points from the mid-vein curve and silhouette curves with a fixed interval. The mesh showed in Fig.3 (a) is generated from Fig.1 (b) by using Delaunay triangulation scheme.



(a) Delaunay triangulation mesh (b) Subdivision mesh

Fig. 3. Delaunay triangulation mesh and subdivision mesh

The initial mesh of a leaf surface generated by using Delaunay triangulation is generally irregular and rough. It is necessary to refine the mesh for later deformation. Currently, we use a simple method to subdivide the initial mesh, which provides several parameters to satisfy requirements for users' interaction. Fig.3 (b) illustrates a mesh subdividing the initial mesh shows in Fig.3 (a).

3.3 Leaf Surface Deformation

We have detailed the mechanism for controlling movement of the venation skeleton in our approach and methods for constructing the leaf mesh. The last step is to deform

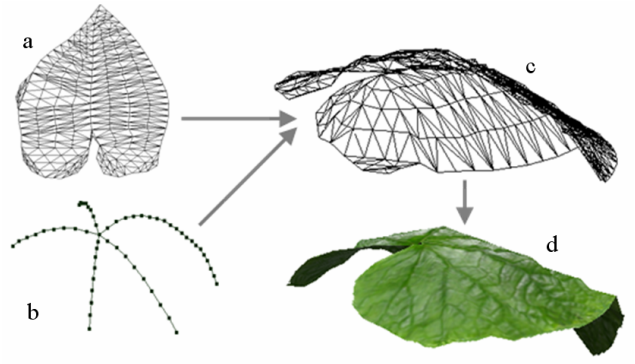


Fig. 4. Deforming leaf surface based on venation skeleton

the leaf surface basing on the deformed venation skeleton. This process can be illustrated as Fig. 4.

Firstly all the vertices in the subdivision mesh of the leaf are banded to the initial venation skeleton. The banding is based on the distance of each vertex to the venation skeleton. Then the initial venation skeleton is deformed by using the method described in Section 4.1. For example, we can generate the shape of the venation skeleton showed in Fig. 4(b) from the initial skeleton showed in Fig.1(c) (with different number of joints in each vein). Lastly the position of each vertex in the mesh is recalculated according to the new coordinate of its banded vertex in the venation skeleton. Fig. 4c displays the resulted mesh. Fig. 4(d) demonstrates the rendering result. The texture mapping is calculated before the deformation.

It needs to be noted that the number of joints in each vein in the venation skeleton will influence greatly the effects of deformation. The larger the number of joints is, the smoother the deformed surface will be. And large deformation needs large number of joints. But larger number of joints also means more computation and more difficult controlling. Users can obtain a satisfactory result by interactive experiment.

3.4 Constrains and Collision Detection

Constraints and collision detection are usually the common issues in surface deformation. For constraint, we have stated that each vertex in the venation skeleton can rotate around a fixed vector. Additionally, the rotation needs to satisfy some extra constraints. In simulating the effect of wilting leaf surface, a vertex in the leaf mesh could not be rotated after it had reached the maximum drooped distance. When simulating curling of a leaf, it needs to avoid overlap of the leaf surface. This can be done by keeping the included angle of two adjacent line sections on each vein being larger than a pre-defined angle.

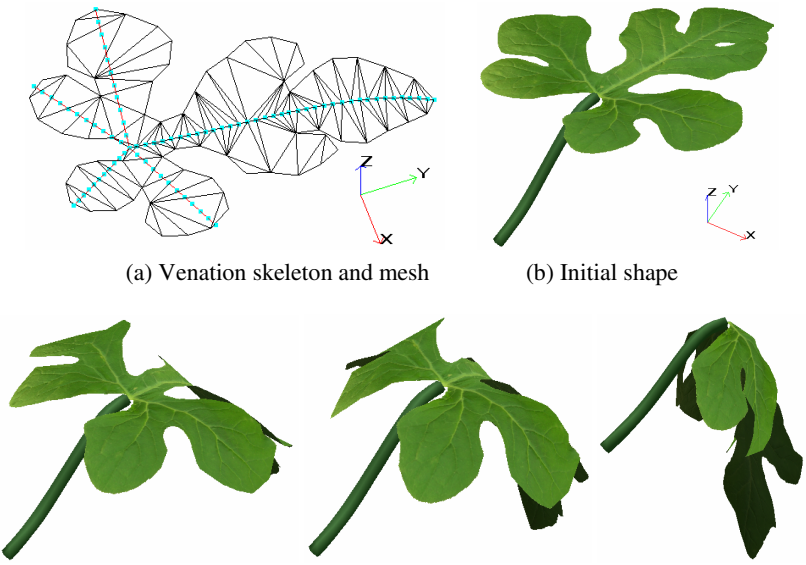
For collision detection, currently we just consider a collision detection to avoid self-Intersect in the deformation of a leaf mesh. During deforming a leaf mesh, each handling currently vertex needs to be checked if its movement will pierce some triangle in the mesh. If no piercing occurs, no response is calculated. If there is an intersection, then we calculate a maximum displacement from the pre-calculated displacement for

the vertex to move while avoid intersection, and correct the displacement of corresponding vertex in the venation skeleton.

4 Applications and Discussion

We implemented our algorithm for venation skeleton-driven leaf surface deformation in C++ on a PC with a 2.8 GHz Pentium D processor and a NVIDIA GeFore 7900 GS graphics card, and use OpenGL to render the results. In this section we report the modeling results.

Firstly we simulate wilting effect of a watermelon leaf, which is a typical lobed leaf. We use a venation skeleton showed in Fig. 5 (a) to control its deformation of the leaf blade. The initial shape of the leaf is showed in Fig. 5 (b), while (c), (d) and (e) demonstrate three wilting effects respectively. We do not apply subdivision to the mesh of the leaf surface, but the results seem are plausible.



(a) Venation skeleton and mesh (b) Initial shape

(c) Slight wilting (d) Medium wilting (e) Acute wilting

Fig. 5. A skeleton structure for watermelon leaf and modeling effects

The second application example is applied to simulate the process of wilting of a cucumber. Fig. 6 demonstrates the simulated results, in which (a) is the initial shape, (b) simulates the slight wilting effect, and (c) is the acute wilting. We use three instances of leaf surface in the cucumber model, and the venation skeleton in each instance is different from each other. The venation skeleton is deformed automatically by rotating the vertices in the skeleton downward from the boundary to the leaf root, by using equation (1), and the leaves above start wilting later than the leaves below do, whereas the speed of wilting can be adjusted by modifying parameter t. The deformation is done in real time, and animation of the wilting process is smooth.

(a) None wilting (b) Slight wilting (c) Acute wilting

Fig. 6. Three stages of a wilting cucumber plant

The application examples above demonstrate that the proposed venation skeleton-driven approach for simulating wilting of leaf surface is effective and flexible, it can generates realistic effects of wilted leaves similar to natural shape. Currently generating the venation skeleton is manual and interactive in our framework, and controlling the motions of leaves in the scale of a plant is still simple. In fact, wilting of leaves are natural response for plant to adapt themselves to the environment basing on their inner state. An attractive area for future work might involve combining our dynamic modeling technique with physiological model of the leaf. In addition, we just consider a plant or a leaf in our framework. It is desirable to simulate the motions of plant leaves in an ecosystem scale.

5 Conclusion

We have presented a model for modeling wilted leaf surface. This model deforms a leaf surface by driving a venation skeleton which is embedded into the geometric mesh of a leaf. The venation skeleton can be created from any polygonal mesh of leaf surface, whereas the polygonal mesh can be captured from real leaves, which makes it easy to create highly realistic leaf appearance models. Furthermore, our model provides an approximately kinematic model of plant leaves for simulating subtle motions of plants.

It needs to be noted that motions of plant leaves would result from a series of complex reasons, as such the mechanism of motions of leaves is not easy to model. The leaf deformation model presented in this paper is an example of a model that provides

intuitive control for the simulating of some motions of plant leaves. An exciting area for future work is the development of a framework for virtual agronomic experiment for broader classes of plants.

Acknowledgments. This work is supported by National High Tech R&D Program of China under grant No. 2007AA10Z226, Beijing Natural Science Foundation of China under grant No. 4081001, and the National 11th Five-year Plan for Science & Technology of China under grant No. 2006BAD10A07.

References

1. Prusinkiewicz, P., Mündermann, L., Karwowski, R., Lane, B.: The Use of Positional Information in the Modeling of Plants. In: Proceedings of SIGGRAPH, Miami Beach, FL, USA, pp. 289–300 (2001)
2. Mündermann, L., Macmurchy, P., Pivovarov, J., Prusinkiewicz, P.: Modeling Lobed Leaves. In: Proceedings of Computer Graphics International, CGI (ed.) Tokyo, Japan, pp. 60–65. CGI (2003)
3. Sung, M.H., Simpson, B., Gladimir, V.G.B.: Interactive Venation-Based Leaf Shape Modeling. Computer Animation and Virtual Worlds 16(3-4), 415–427 (2005)
4. Nath, U., Crawford, B.C.W., Carpenter, R., Coen, E.: Genetic Control of Surface Curvature. Science 299, 1404–1407 (2003)
5. Sharon, E., Roman, B., Harry, L.: Geometrically Driven Wrinkling Observed in Free Plastic Sheets and Leaves. Physical Review E 75, 1–7 (2007)
6. Beaudoin, J., Keyser, J.: Simulation Levels of Detail for Plant Motion. In: ACM SIGGRAPH/Eurographics, Symposium on Computer Animation, pp. 297–304 (2004)
7. Jirasek, C., Prusinkiewicz, P.: A Biomechanical Model of Branch Shape in Plants. In: Lantin, M. (ed.) Proceedings of the western computer graphics symposium, Whistler, Canada, pp. 23–26 (1998)
8. Hart, J., Baker, B., Michaelraj, J.: Structural Simulation of Tree Growth and Response. The Visual Computer 19, 151–163 (2003)
9. Wang, I., Wan, J., Baranoski, G.: Physically-Based Simulation of Plant Leaf Growth. Computer Animation and Virtual Worlds 15, 237–244 (2004)
10. Amenta, N., Bern, M., Kamvysselis, M.: A New Voronoi-Based Surface Reconstruction Algorithm. In: Proceedings of ACM SIGGRAPH, Orlando, Florida, USA, pp. 415–421 (1998)

The Technical Research and System Realization of 3D Garment Fitting System Based on Improved Collision-Check Algorithm

Qingqing Chen, Junfeng Yao, Hanhui Zhang, and Kunhui Lin

Software Department of Xiamen University, Xiamen, China, 361005

Abstract. Nowadays, the study of 3d garment fitting technology is going mature. There has a series of research results in clothes-stitching, collision-checking and so on. But the trad collision-checking algorithm could result in endless loop, and it is more complex. There is no internal system which is perfect and has been spread. This system improved the trad constructing process of bounding-box, avoiding the endless loop that appeared in trad bounding-box constructing process. The system used easier algorithm to check collisions, reduce the complexity and can do precise and fast collision-check between clothes and human body. The 3D garment fitting system introduced in this paper can simulate the virtual fitting of 3D garments and set out the effect of it. This system uses Client/Server mode. Users download Client and garment patches and run them at local computers. In this way, the system can solve the slow-updating problem of 3D images in 3D internet fitting system generated by slow net rate. It uses imported the 3DS 's mannequin to setup human body models, and solves the singular problem of motion and pose and the lacking problem of face and extremities generated by human platform modeling. Users design 2D garment patches, then the system would realize the transforming from 2D garment patches to 3D garments automatically. So the system can update garment style rapidly. It also has dynamic simulation function and realizes the reality of different fabric. We trust that this system could excite the potential of sell effectively after it is perfected and is extensively used.

Keywords: Spring-Mass model; collision-check; 3D garment fitting system.

Introduction

3D garment fitting system is the result of the science's development. The way to buy clothes is changing while the science technology and the internet are developing. At first, buyers needed to select clothes by fitting it on body in shops, then 2D level's garment fitting systems appeared which used picture-paste fashion to simulate the effect of fitting. Now, various 3D garment fitting systems have appeared at home and abroad. All of these represent the significance of technology in human life.

3D garment fitting system includes the human body modeling, the 3D garment modeling and the simulation function. 3D garment modeling includes two phases: first,

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 736–744, 2008.

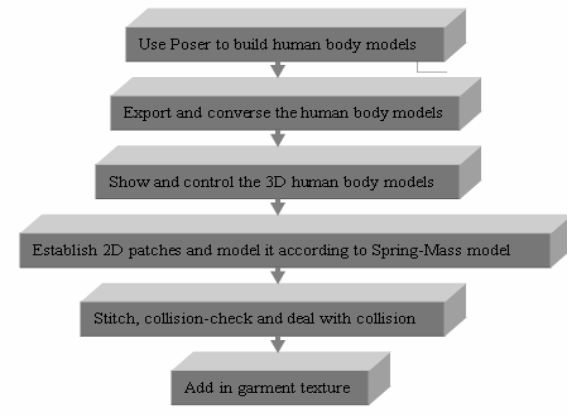


Fig. 1. The main processes of the system's constructing

generating the initial 3D garments according to 2D patches, which means completing the transition from 2D to 3D; second, adding texture into the initial 3D garments and simulating drape and pendency to achieve real effect of fitting. This paper introduces the developing process of the whole system and the technology means the system used.

1 The Developing Process of the System

Whether we can Real-timely and efficiently control the internet garment fitting system lies on the rate of the internet. It means that the increasing of the internet rate is important. If the internet delay always happens, the Real-time effect of the fitting would not be acquired and we would waste a lot of time in the waiting of updating 3D pictures. As a result, we use the C/S mode for the system. Users only need to download Client and garment patches and run them at local computers. In this way, it can solve the problems generated by slow internet rate.

2 Main Processes

2.1 Human Body Modeling

There are two main ways to establish human body models: building models based on measure-rebuild method and building models based on imported the 3DS 's mannequin.

Model-building based on measure-rebuild method: the best strongpoint of it is the accuracy and the controllability of the human body size. But there also exists shortages. It's ok to use human platform instead of human body models in the design of the garment structure. But human platform can't replace human body models in some areas because of the singular problem of motion and pose and the lacking problem of face and extremities generated by human platform modeling. Such as the stage-exhibition of

the garments in virtual reality and the dynamic garment fitting function. In this paper, we design an interface faced to 3DS files to import human body models in 3DS format[2][3][4][5]. Imported the 3DS's mannequin translates the human body models established by Poser and saves them as 3DS files, then it analyzes their structure, imports them, and builds the human body models by rebuilding the 3D human bodies according to the information of the 3D coordinates.

(I) We can design two classes to read 3DS files according to the analyzing of the "block" structures of 3DS files[7]. The two classes can be described as Fig. 2.

(II) After importing the 3DS files, we can build human body models by using lots of triangle planes according to the vertexes and the normal vectors read.

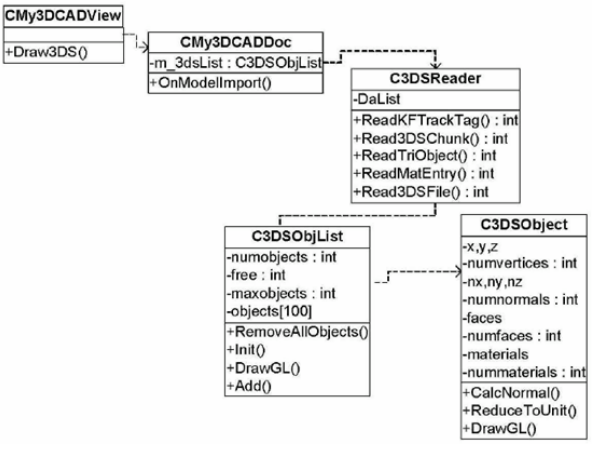


Fig. 2. The UML class diagram for importing 3DS models

2.2 Garment Modeling

The mapping process from 2D patches to 3D garments is a complex and flexible deformation process. It needs to meet the following conditions[8]: first, it should keep the area of the patches during the mapping process; second, patches between each other should satisfy the correct relations; third, no collision happens during the mapping process. The frequently used modeling method based on physics. It contains Flexible-Distortion model, Particle-System model, Finite Elements method, Spring-Mass model and so on. After analyzing we found that the Spring-Mass model is simpler and it can acquire more real simulation effect and has higher simulation speed.

The fabric deformation model we built is based on the Spring-Mass model built by X Provot. 2D patches and 3D garments are all dispersed and are expressed as Spring-Mass system composed by regular triangle grids. The vertexes of the grids are particles. The sides of the grids are springs. Each particle connects with particles around it by springs. The relationship between particles is the stretching effect. According to the mechanism capability of the fabric, the springs can be divided into

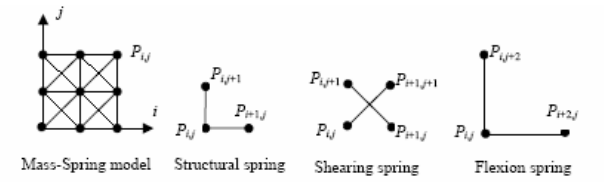


Fig. 3. Mass-Spring model and 3 types of springs

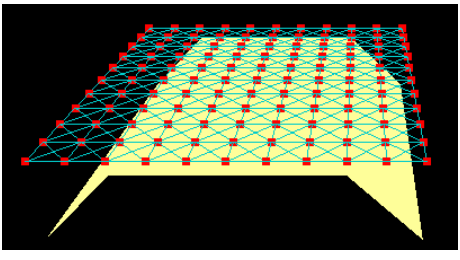


Fig. 4. The patch constructed by 3types of springs

three types: structural spring, shearing spring and flexion spring[9]. They can be described as Fig.3. The patch could be constructed by these 3 types of springs as described by Fig.4.

2.3 Patches-Stitching Process

(I) Import the patches designed by CAD for 2D garments.

(II) Select the corresponding sides of the patches that need to be stitched.

(III) Disperse 2D patches and form the initial Spring-Mass system.

First, disperse the patches and form regular quadrangle grids. Connect the diagonals of the grids to form regular triangles, thereby a Spring-Mass system is established. The vertexes of the triangles are particles and the sides are the corresponding springs.

Second, add in various springs according to the relationship between particles.

(IV) Put the patches alternatively at the initial location near the human body models.

(V) Compute the deformation model dynamically.

We add in a stitching power on the corresponding sides of the patches according to the stitching information of the patches. Under the effect of the power, the gravity and the internal spring between particles, the 2D patches would be deformed gradually and would be stitched with each other. The whole process of the stitching is a dynamic and iterant process. It would check that if any collision happens and deal with it when it happens.

2.4 Collision-Check

The collision contains the fabric collision and the collision between parts of the clothes. It means that self-collision-check has two parts[10]. The usual 3D models (contain the

fabric and the objects around it) are expressed as triangle grids. In this way, the process of the collision-check is just the way to check if any penetration happens between particles and triangles and if it happens between sides of triangles. But, if we check each particle-triangle pair or we check each pair of sides, the capacity of the computation is too bulky. In order to reduce the bulkiness, we use a method based on AABB hierarchical bounding volumes[11] to exclude the pairs of particle-triangle and the pairs of sides that would not intersect with each other. We built a AABB tree for human body models(as Fig.4 describes) and patches and then we just check the pairs of particle-triangle and pairs of sides that intersect in hierarchical bounding volumes.

2.4.1 Constructing Bounding-Boxes

There has two ways to construct AABB tree: top-down and bottom-up. We use top-down to construct it (this can be described by Fig.5). The algorithm is:

(1) Working out the coordinate of vertexes included in the root V.

(2) Working out the AABB of the node V.

(3) Dividing the AABB into two subsets along the longest axis of it according to the cg of the triangles included in this AABB.

(4) Using the two subsets as two roots. If every AABB is a leaf, finish the constructing process, else, return to step (2).

The AABB tree we got is a complete binary tree, and each leaf node contains only one triangle.

If we restrict that each leaf node only contains one triangle according to the trad way, the bounding-box constructing process will result in endless loop when the following situation happens (Fig.6).

If this situation happens, the cg of three triangles included in A would be in the same subset when we dividing the bounding-box A. This would lead to the following results:

(1) The bounding-box of A.children[0] is the same as the bounding-box of A.

(2) The bounding-box of A.children[1] is null.

In order to avoid the endless loop, each leaf node should be set to include one more triangle, and we used the following method:

(1) If the face number of A's subsets is the same as A's, then we didn't do the recursion for A;

(2) We did collision-checking according to the actual triangle number included in each leaf node.

After the constructing of bounding-box, we exclude the object-pairs that would be impossible to interact with each other according to the intersect-checking of the bounding-box of objects.

2.4.2 Collision Checking

The precise collision-check includes two steps:

(1) Check whether the vertex be colliding with the plane where certain triangle is located.

(2) Check whether the colliding-node is inside the triangle.

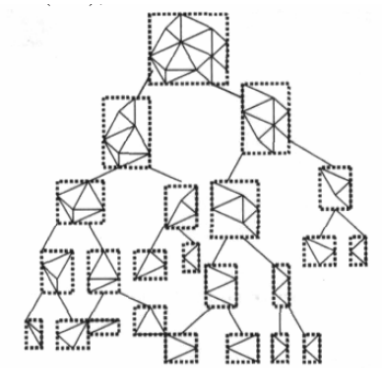


Fig. 5. The hierarchy of AABB tree for human body models

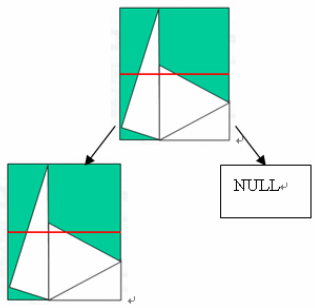


Fig. 6. The endless loop

The trad collision-check algorithm is more complex. Here we used easier algorithm[11] (as Fig.7describes):

(1) The originate position of the vertex a is p0, the new position after time step t is p1, p is the normal vector of the plane that might be collided with a. Find out the length of vector $\overrightarrow{pop1}$ in the direction of p.

(2) A is the random point in the plane. Find out the length of vector $\overrightarrow{Ap1}$ in the direction of p. Mark the value h'.

(3) If the value of h'/h (mark the value fPercentage) is between 0 and 1, vector $\overrightarrow{pop1}$ would collide with the plane certainly. We could work out the coordinate of the cross point M by using the formula

$$p0 + |p0p1| * (1\text{-}fPercentage). \tag{1}$$

(4) If the sum of $\angle AMC, \angle CMB, \angle BMA$ is 360°, then the cross point M is surely inside the triangle ABC. We could make conclusion that the vector $\overrightarrow{pop1}$ is colliding with the triangle ABC.

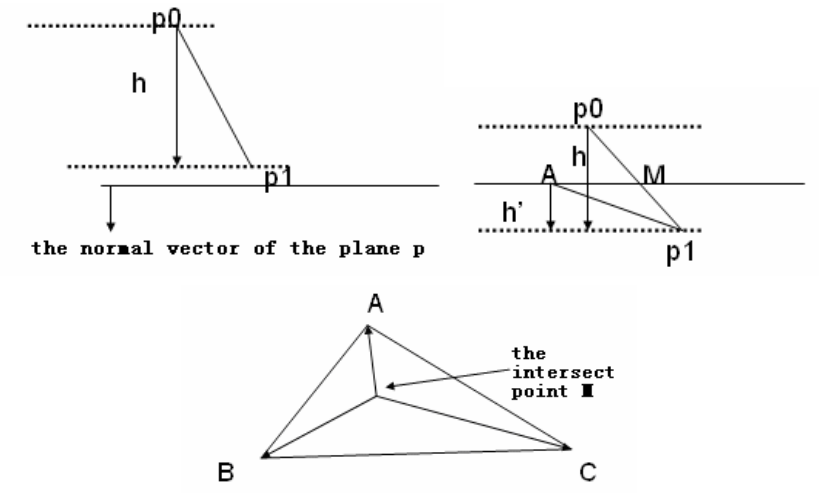


Fig. 7. The easier collision-checking algorithm

On the other hand, we use Provot's[12] triangle-curvature arithmetic to predigest the computation. When the angle between normal vectors of neighboring triangles is small, collision would not happen. Only when the angle excesses the key value does collision happens. The system can exclude most circumstances in which the triangles would not intersect with each other and predigest the computation by computing the surface curvature of neighboring triangles. The described method reduces the times of the collision-check between human body models and patches that wouldn't interact with each other. It uses $SA_{T[13]}$ to judge the overlapping between bounding volumes to reduce the arithmetical complexity and increase the arithmetical efficiency.

The effect of the collision between the cloth and sphere and the collision between cloth and round table could be described as Fig.8 Fig.9and Fig.10

Fig. 8. The collision between sphere and cloth-1

Fig. 9. The collision between sphere and cloth-2

Fig. 10. The collision between cloth and round table

3 Conclusion

The purpose of the 3D garment fitting system is to realize the simulation for the 3D human body models and the 3D garments. The main processes of it are 3D human body modeling, changing the plane structure from 2D patches to the solid structure of 3D garments and the simulation for the garment fitting according to the material capability of the fabric. This paper discusses the key technology of the fitting system's constructing in detail. It contains 3D human body modeling and the importing, the mapping model of 2D patches and 3D garments, the improved collision-checking technology and responding process. The 3D garment fitting system described in this paper has the characteristics of rapid updating rate and real simulation effect. It can realize more complex garment style. We trust that this system could excite the potential of sell effectively after it is perfected and is extensively used.

Acknowledgement

Supported by Program for New Century Excellent Talents in Fujian Province University and the 2006 Foundation of Department of Science & Technology in Fujian Provincial (2006H0035).

References

[1] McCarney, J., Hinds, B.K., Seow, B.I., Gong, D.: Dedicated 3D CAD for garment modeling. Journal of Materials Processing Technology 107, 31–36 (2000)

[2] Dongmei, Y., Shengyuan, Z., Weicheng, L.: The integration of OpenGL and 3D Studio MAX to realize 3D simulation. Application Technology (2), 33–35 (2004)

[3] Bin, F.: The application of 3DMAX models in OpenGL. Journal of GuiZhou Industrial University (6), 45–49 (1999)

[4] Wenguang, Z., Zhongxue, L., Cuiping, L.: The Visualization simulation of VC++ faced to engineering,The integrative technology of OpenGL and 3DS. Journal of BeiJing Technology University (6), 53–565 (2001)

[5] Zhen, Z.: Poser4 3D Cartoon Design-God of War, vol. 1. National defense industry press, Beijing (2004)

[6] Chen, S.F., Hu, J.L., Teng, J.G.: A finite-volume method for contact drape simulation of woven fabrics and garments. Finite Elements in Analysis and Design 37, 513–531 (2001)

[7] Peng, W.: Studies on Constructing Mannequin and Fashion Pattern Design Technology in Virtual Reality. Master's degree thesis of JiangNan University, 35

[8] Fan, J., Zhou, J., Wang, Q.F., Yuan, M.H.: 2D/3D isometric transformation using spring-mass system. Journal of Software[J] (in Chinese with English Abstract) 10(2), 140–148 (1999)

[9] Hui, L., Chun, C., Bole, S.: Simulation of 3D Garment Based on Improved Spring-Mass Model. Journal of Software[J] (14), 620–621 (2003)

[10] Changfeng, L., Yi, X.: Cloth 3D Real-Time Simulation. Journal of Computer-Aided design & Computer graphics[C] (18), 1375 (2006)

[11] Chengying, G., Ning, L., Xiaonan, L.: The establishment of fabric model and the dealing with conllision-check in virtual fitting. Computer Applications (5), 34–37 (2002)

[12] Provot, X.: Collision and self-collision handling in cloth model dedicated to design garments[C]. In: Proceedings of Graphics Interface 1997, Kelowna, pp. 177–189 (1997)

[13] Louchet, J., Provot, X., et al.: Evolutionary identification of cloth animation models[A]. In: Terzopoulos, D., Thalmann, D. (eds.) Proceedings of the Computer Animation and Simulationc 1995[C], pp. 44–54. Springer, New York (1995)

[14] Hongyan, S., Jun, L.: The developing trend of garment internet shopping. Chemical Fiber&Textile Technology (3) (2005)

[15] Kang, Y.M., Choi, J.H., Cho, H.G.: Fast and stable animation of cloth with an approximated implicit method. In: Thalmann, N.M. (ed.) Proceedings of the Computer Graphics International 2000, pp. 247–255. IEEE Computer Society Press, Los Alamitos (2000)

[16] Breen, D.E.: Computer graphics in textiles and apparel modeling. IEEE Computer Graphics and Applications 16(5), 26–27 (1996)

[17] Ng, H.N., Grimsdale, R.L.: Computer graphics techniques for modeling cloth. IEEE Computer Graphics and Applications 16(5), 28–41 (1996)

[18] Hadap, S., Bangerter, E., Volino, P., Thalmann, N.M.: Animating wrinkles on clothes. In: Ebert, D.S., Gross, M., Hamann, B. (eds.) Proceedings of the IEEE Visualization 1999, pp. 175–182. IEEE Computer Society Press, Los Alamitos (1999)

[19] Volino, P., Thalmann, N.M.: Implementing fast cloth simulation with collision response. In: Thalmann, N.M. (ed.) Proceedings of the Computer Graphics International, pp. 257–266. IEEE Computer Society Press, Los Alamitos (2000)

[20] Okabe, H., Imaoka, H., Tomiha, T., Niwaya, H.: Three dimensional apparel CAD system. Computer Graphics 26(2), 105–110 (1992)

Reconstruction of Tree Crown Shape from Scanned Data

Chao Zhu[1,2], Xiaopeng Zhang[1,2,*], Baogang Hu[1,2], and Marc Jaeger[3]

[1] Sino-French Laboratory LIAMA, CAS Institute of Automation, Beijing, China
[2] National Laboratory of Pattern Recognition, CAS Institute of Automation, Beijing, China
[3] INRIA-Saclay, Project DigiPlante, CIRAD AMAP, Montpellier, France
{czhu, xpzhang, hubg}@nlpr.ia.ac.cn, jaeger@cirad.fr

Abstract. Reconstruction of a real tree from scattered scanned points is a new challenge in virtual reality. Although many progresses are made on main branch structures and overall shape of a tree, reconstructions are still not satisfactory in terms of silhouette and details. We do think that 3D reconstruction of the tree crown shapes may help to constrain accurate reconstruction of complete real tree geometry. We propose here a novel approach for tree crown reconstruction based on an improvement of alpha shape modeling, where the data are points unevenly distributed in a volume rather than on a surface only. The result is an extracted silhouette mesh model, a concave closure of the input data. We suggest an appropriate scope of proper alpha values, so that the reconstruction of the silhouette mesh is a valid manifold surface. Experimental results show that our technique works well in extracting the crown shapes of real trees.

Keywords: tree crowns, reconstruction, Delaunay triangulation, alpha shape.

1 Introduction

With the current development of virtual environment establishment, product design, digital entertainment, antique protection, and city programming 3D geometry model construction and processing is now an active development area. 3D geometry modeling is regarded as the fourth digital multimedia in addition to digital audio, digital image, and digital video. 3D geometry models are normally used to represent object surface to identify extendedly shape and appearance attributes.

With the advancement of 3D scanning technology, more and more 3D digital scanners are popularly used for different applications. Rich details of the object shape can be acquired from scanned data with dense sampling points (point cloud), where no topological connection relations are included. It becomes important to develop new processing methods to represent, to process, to reconstruct and to render these highly complex geometric bodies. Reconstruction of geometry model is one of the important research topics in modern virtual reality.

Trees are typical objects in virtual reality, so it is very important to reconstruct and to represent the real trees. Tree reconstruction can be used in many applications, including digitization of vegetation scenes, design of a new scene, digital entertainment, and so on.

* Corresponding author.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 745–756, 2008.

Reconstruction of the tree crown shape is useful to model a real tree and in various research fields interested by growth simulation of virtual trees for light interception, biomass evaluation, and so on. Many researches have been carried on surface reconstructions, but the shape of a tree crown is more difficult than that of a usual solid object in its heavy occlusions and its high complexity in geometry and topology.

Reconstruction of a real tree crown from scattered scanned points is a new challenge in virtual reality. The points are unstructured, unevenly distributed, and sampled from non-manifold shapes; it is very difficult to define typical boundary points in a tree crown in accordance with the visual perception. Rich concavity is another feature of the crown shape of a real tree. Classical surface reconstruction techniques do not work for such tree crown data. Other difficulties are that the data have no topological information among points, their 3D distribution is not even at all, and they are not complete; so it becomes rather difficult to reconstruct it with sufficient details.

Alpha shape is a new technique [1] in the classification of all the simplexes from 3D Delaunay triangulation of a 3D point set, and the result of this classification is three categories: the internal simplexes of the shape, the regular one and the external one. With a proper heuristic alpha value specified by the user [1], a concave silhouette shape of a point set sampled from a regular manifold surface can be constructed.

We will improve this approach to the point cloud data acquired from the scan on a real tree. Because of the limitations of alpha shape technique and complexity of scanned tree data, it is not possible to have all details of the tree crown reconstructed to a regular mesh through a direct application of the alpha shape technique [1]. In this paper we solve this problem using a range of alpha values and testing the close property of the constructed mesh, so that the mesh model is a concave closure of the point data.

The structure of this paper is as follows. Related work in shape information analysis and plant modeling is introduced in section 2. Fundamental knowledge of our method is described in section 3. Technical details of this new approach are described in section 4. Experiments of this technique to reconstruct tree crown shapes are shown in section 5. Conclusions about this technique and further investigation are described in section 6.

2 Related Work

In the past decades, many methods have been developed on point shape processing and shape modeling of complex objects including plants, but with unequal results.

2.1 Point Geometry Processing

Jarvis [2] was the first to consider the problem of computing the shape as a generalization of the convex hull of a planar point set. A mathematically definition of the shape was developed by Edelsbrunner in [3]. For 3D points, Boissonnat [4] suggested to use Delaunay triangulation to “sculpture” a single connected shape of a point set.

In the frame of projects such as the digital Michelangelo project [5] at Stanford Computer Graphics Lab in the 2000’s, and with the improvement of computer hardware, a numerous number of research papers have been published on point cloud processing and rendering. Point geometry processing and analysis became an active research topic.

2.2 Plant Modeling on Knowledge and Rules

The different approaches of 3D tree model construction can be roughly classified into three categories: botanical models, geometrical models, and digitized models from real plants.

There are a numerous methods to simulate real plant appearance. Many early methods were based on rule iterations (botanical, physical, geometrical, mathematical), or simply based on strong user control with advanced dedicated patterns. In the 1980s, modeling by botanical rules appeared, and produced nice findings, researchers tried to simulate the growth of natural plants, plants could be constructed by some botanical rules or grammars. AMAP [6] modeling method is based on bud life cycles of botanical knowledge with real measurement data (on plant topology). This modeling method clear reflects the growth mechanism of plants, including space occupation and the location of leaves, fruits, and flowers. L-systems presented by a Lindenmayer and Prusinkiewicz were broadly applied to describe the growth process of plant organs, which were based on fractal pattern [7, 12].

GreenLab [13] modeling approach is put forward as a mathematical model, which simulates interactions of plant structure, leaves, trunk, branch and function. This model can exactly engender the dynamics of plant, architecture and geometry of woody plants, because of internal competition for resources, leaves sizes are different, and growth of pruning can also be simulated.

These methods, used mainly in biology research fields are not dedicated to control the 3D plant shape, and cannot easily do it, but aim to understand plant shape as the result of a dynamic. It as be cited that this kind of model is not suitable to construct a 3D models of real tree by using botanical methods [10].

Geometrically interactive modeling is another way to model virtual plants. Although this method does not strictly follow the botanical rules, but visually realistic trees can be produced [14]. In general, given 3D skeleton points of real plant, 3D model of each branch can be generated with generalized circular cylinders [15]. Prism model is a simplified application of this method. This approach is widely applied in some plant software such as Xfrog, if combining rule-based method with traditional geometric modeling approach. Nice 3D plant model could be produced, such as flowers, bushes, and trees [8, 14].

To summarize, these rule based or pattern based methods used to build the real plant faithful to botanic knowledge or appearance, can produce visually very realistic plants, although they could not be used to model a specific existing real plant.

2.3 Digitalization of Real Plants

New modeling methods have been developed to digitalize real plants in very recent years [9-11]. These methods can be used to reconstruct the trunk, the branches, and the leaves, but the realism of the reconstructed model is still different from the real shape due to the lack of crown silhouette shape information.

Plant digitization aims to reconstruct the shape of real plants from the information digital instruments. The most popular techniques are the use of 3D laser scanner [9] or the use of digital photos [10].

When scanning a real plant from a single viewpoint many occlusions occur. In particular, many leaves do usually hide branches and other organs from the view. One way to make reconstruction efficient is to work both on plant branching structure reconstruction and on plant crown reconstruction. The idea of the proposed approach is to consider that when processing the branch reconstruction of a real plant, we must constrain the silhouette of branches from the crown shape.

By scanning a real tree, we have a point cloud data set, from which we could reconstruct the shape of the real tree crown by combining existing methods.

Considering the branch structure, we may underline several interesting works. Cheng [16] reconstructs a real tree from a range image, using generalized circular cylinders to fit incomplete data and compute the skeleton based on axis direction. Pfeifer [17] introduces a set of algorithms for automatically fitting and tracking cylinders along branches and reconstructing the entire tree.

With the appearance of advanced precise digital camera and laser scanner, the development of digital plant is accelerated. Image-based and laser-scanning based methods have come up to produce 3D model of real trees in nature. Shlyakhter [18] builds a 3D model of tree from a set of photographs. His method constructs the visual hull of tree first, then a plausible skeleton is built up from medial axis of visual hull, and finally L-system is applied to construct branches and leaves. Teng [19] reconstruct 3D trunk of plant only from two images, this method only estimates skeleton and radius of branches roughly. Quan [10] also models a plant from digital image. Their work focus on reconstruction of big leaves, branches are reconstructed by interaction.

These image-based approaches can build 3D plant from images of different viewpoints, but because of inevitable noise of images and error of camera parameters, the accuracy of those methods is limited.

The approaches of Xu [9, 11] are based on some prior knowledge. A skeleton is first constructed by connection of the centroids of points, which have an analogous length of the shortest path to a root point. Then the corresponding radius of skeleton nodes could be computed by the allometric theory. Leaves are constructed in the end, so that the reconstructed tree is visually impressive.

However during the reconstruction, the methods of imaged-based or 3D laser scanning data based first construct the skeleton, and then construct leaves, but because of much occlusion, the reconstructed skeleton is incomplete. 3D laser scanner could not scan the thin branches because of its limited precision.

But we have to reconstruct these thin branches for the architecture shape of real plants in botany and digital forestry and for high visual impression in virtual reality.

We must thus construct the shape of tree crown to constrain the reconstruction of thin branches.

3 Algorithm Bases: Alpha Shape

Alpha shape was proposed in 2D by Edelsbrunner [3], and was then extended to 3D in [1]. This method can be used to reconstruct object surface from an unorganized point cloud. Our reconstruction of concave tree crown is based on this technique.

3.1 Delaunay Triangulation

A set P of points can be used to construct a complex if the points do not lay in a plane. Delaunay triangulation is a natural choice to do it. In literature, different Delaunay triangulation techniques are proposed [20-22], where Lawson flip method is a typical one. in Lawson's method, the tetrahedron bounding the point set P is constructed at first, and the other points are inserted into the triangulation one by one then. Each time, the triangulation is optimized to satisfy the Delaunay property: the circumsphere of every tetrahedron does not contain any other points. Those tetrahedrons, which do not satisfy a local Delaunay property, are flipped.

The flip process in 3D can be described as follows. The triangulation in 3D is a set of tetrahedrons constructing a simplical complex. We will explain the case of two tetrahedrons incident to a triangle ace (Figure 1). If the circumsphere of tetrahedron $aecd$ does not contain b and circumsphere of tetrahedron $aecb$ does not contain d, it can be said that (the triangle) aec is local Delaunay. Otherwise, this situation can be modified inserting a new edge bd inserted. Therefore the complex is a Delaunay triangulation.

The result of Delaunay triangulation of the point set is its convex hull composing several tetrahedrons.

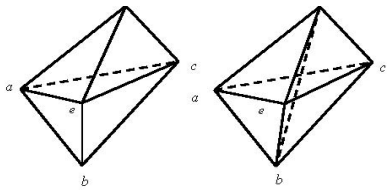


Fig. 1. Flipping in three dimensions

3.2 Alpha Shape

The concept of alpha-shapes formalizes the intuitive notion of *shape* for spatial point sets on user's selection. Alpha-shape is a mathematically well-defined generalization of the convex hull. Its result is a series of subgraphs of the Delaunay triangulation, depending on different alpha values. Given a finite point set, a family of simplexes can be computed from the Delaunay triangulation of the point set, a real parameter alpha controls the desired level of detail. All real alpha values lead to a whole family of shapes. The alpha-shape of a point set is made up of the set of points, edges, triangles and tetrahedrons, which satisfy the constraint condition: the alpha test [1]. This test applies for each a triangle t of the triangulation. If t is not on the boundary of the convex hull, there must be two tetrahedrons p,q, which are incident to t. Tetrahedrons p and q are tested to be in the circumsphere of t or not. If they both are not in that circumsphere, and the radius of the circumsphere is less than the alpha value, t is said to satisfy alpha test, and it is regarded as one member of the alpha shape. So alpha-shape is a subset of the triangulation.

If we let alpha be large enough, the shape is the convex hull of the points set. If alpha approaches 0, no tetrahedral, triangles and edges could pass the alpha test, so the alpha

shape is the points set. With the adjustment of the alpha values, this subset can follow the topology of the points set. So, if we choose a proper value for alpha, we will find a reasonable surface for a tree crown.

The alpha shape is a sub-complex of the Delaunay triangulation of the points set P. This can be explained in the following. There is a ball eraser with alpha as its radius, and it could move to all possible positions in the 3D space and with no point of P included. This eraser will delete all simplexes whose size is bigger than alpha and it can pass through. So the remaining simplexes construct the alpha shape.

4 Shape Construction of Tree Crown

The most impressive aspect of a tree is the silhouette of its crown, so the shape of the crown surface is one important aspect for tree reconstruction for the virtual environments. We can only acquire discrete points of the crown with the most recent sensors in nowadays. Normally the data with 3D laser scanner are range images, each of which is obtained from the scan at single viewpoint.

Point cloud from leaves determines the shape of tree crown. Since branches support leaves in the architecture, branch reconstruction is important also. If we do not have the branch model, we do not know how to locate the leaves. Reconstruction of branches consistent to tree crown should be the main target of the reconstruction of a real tree. It is very hard to reconstruct tree branches directly since shape information of the point data is rather weak. The data for branches are incomplete due to the occlusion of leaves and other branches. On the other hand, some little twigs cannot be scanned because of precision limit laser spots.

If we build up the surface of a tree crown from the scanned data, the reconstruction of tree branches will be easier under the control of tree crown surface. Otherwise, the reconstruction result might be different from the real tree, so not faithful to be applied to tree measurement.

4.1 Analysis of Scanned Data of a Real Tree

It is an ordinary technique to sample the surface of real object using 3D laser scanner, and then to reconstruct the shape from the sample data with limited precision. This point cloud data describe the geometry and the appearance attribute of objects surface. The normal point cloud is densely sampled from continuous or smooth surface, although the data is unorganized and irregular. A number of successful methods have been presented to deal with these data and to reconstruct appearance of object.

Plants, such as trees, have too many organs and its structure is too complex. A tree is made up of trunk, branches, and a huge number leaves. The point cloud data of tree is not sampled from a manifold surface, so it is more irregular than those from other data from the manifold surface. The points from leaves are even more irregular. The density variation of point cloud from leaves may be very large. Thus, traditional technique does not work for these objects. Special methods should be developed to reconstruct real plants. In order to keep the shape of plants, branch skeleton extraction and construction of plant crown should be included. One difficulty of this work is that the points from

branches and those from leaves are mixed together, so that it is hard to initialize the work of shape analysis.

4.2 Building the Mesh Model of Tree Crown

From the above analysis and the range image data acquired from a single scan in Figure 3(a) and Figure 4(a), we can recognize the dense region and the sparse region of the data by observation, but this recognition process is very difficulty to be performed in a computer. The points from the tree side facing the scanner and the region with dense leaves (the side of a tree facing the sun, for example) are denser. There may be some interstices among dense leaves. When we scan a tree, laser lights will pass the interstice and meet the branch or leaves at another side of tree, or pass through the tree. So there should be holes in the data. Although we can distinguish the dense region, the sparse region, the convex region and concave region, the algorithms processing very dense point set will make mistakes in topological reconstruction. Therefore, we must construct topological structure of points at first, where Delaunay triangulation is an ideal choice.

Our algorithm contains four steps:

The first step is to triangulate point set $P=\{p_i\}$ with Delaunay triangulation, so that a set of connected tetrahedrons $T=\{T_j\}$ are obtained. Flipping method in [23] is adopted to correct irregular triangulation in $P=\{p_i\}$. All tetrahedrons $T=\{T_j\}$ will constitute a convex solid, the shell of this solid is a convex hull.

The second step is to compute all radii, $R(T_j)$, of circumsphere of every tetrahedron after triangulation. This value will be one attribute of a tetrahedron T_j. The radii, $r(F_k)$, of the circumcircle of each face of a tetrahedron arecomputed also, and they are thought of as an attribute of each face.

The third step is to classify tetrahedrons $\{T_j\}$ and their all faces. The rule to classify all T_j is the size of $R(T_j)$. This classification is performed by the relation of $R(T_j)$ with threshold α and, where α is specified by users. The scope of α should be proper. Then all tetrahedrons are classified into two categories according to a real value α: interior tetrahedrons and exterior tetrahedrons. If $R(T_j)>\alpha$, T_j is classified as an exterior tetrahedron. Otherwise, it is classified as an interior tetrahedron. All faces $\{F_k\}$ from each tetrahedron T_j are classified into three categories also: interior faces, exterior faces and boundary faces. The classification role is as follows. If a face on the convex hull belongs to an exterior tetrahedron, it is an exterior faces; otherwise, if it belongs to an interior tetrahedron, it is a boundary faces. For each face not on the hull, if it is an intersection face of two exterior tetrahedrons, it is an exterior face. If it is an intersection face of two interior tetrahedrons, it is an interior face. If it is an intersection face of one interior tetrahedron and one exterior tetrahedron, it is a boundary face. All boundary faces will construct a mesh, and this mesh M will be an concave approximation of the crown.

Let rmax be the largest radius of all $R(T_j)$ and all $r(F_k)$, and Let rmin be the smallest radius of all $R(T_j)$ and all $r(F_k)$. We acquire an interval $[A,B]$, where $A=\lambda r\min$, $B=\mu r\max$, $\lambda=0.9$, and $\mu=1.1$. The α value should be confined to the

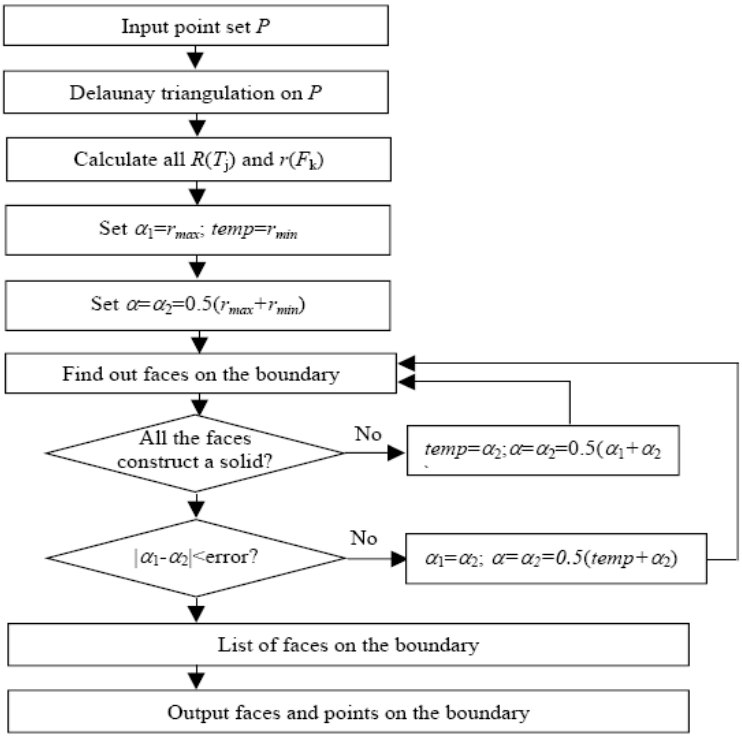


Fig. 2. Pipeline of this algorithm

interval $[A,B]$; otherwise, if $\alpha > B$, the mesh M will be a convex hull, and if $\alpha < A$, the mesh M will not be a solid.

The fourth step is to test the validity of specific alpha values, so that the mesh M builds a boundary surface of a manifold. If the alpha value is set larger than B, boundary points are on the convex hull, so the mesh cannot be concave. If the alpha value is set smaller than A, some sample points are isolated from in the solid, so the reconstructed shape is not complete. Those both extreme cases are not interesting for tree crowns. Therefore, α must lay in interval $[A,B]$. Finding the proper α value is an iterative process. We initialize α as the average value of A and B. In each iteration step, we check if the boundary triangles constitute a manifold surface; if so, the alpha value can be reduced, if not, it is increased.

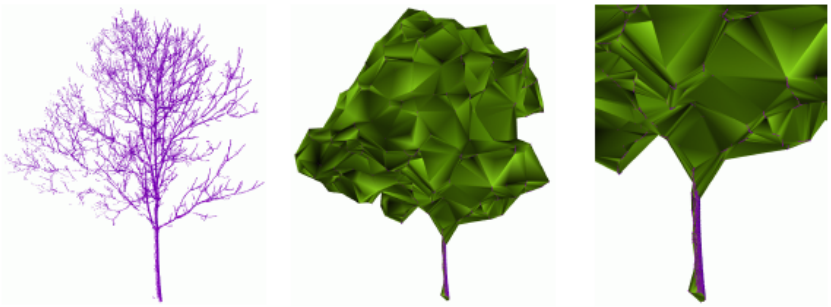
Figure 2 shows the pipeline of this approach.

5 Experiments and Discussion

Our algorithm is written with C Language with the support of OpenGL for graphics. Tests were held on a PC with P4, 3.0GHz processor and 1G RAM. CGAL library is used to perform Delaunay triangulation [24]. Our experimental results of concave tree crowns are shown with local illumination.

(a) (b) (c)

Fig. 3. Extraction of the Crown shape of a Maple tree; (a) is the source point cloud data; (b) is the boundary point cloud data; (c) is the extracted boundary mesh model



(a) (b) (c)

Fig. 4. Extraction of the crown shape of a Candlenut tree; (a) is the source point cloud data displayed with a cube for each point; (b) is a comparison of the extracted boundary mesh with the source point cloud date; (c) is a close view of (b)

We reconstruct the shape of tree crowns with two data sets of two trees. The first one is a single scan of a 20-meter high maple tree with leaves. Figure 3 (a) shows the original point model of the maple tree of 114997 points displayed with a cube for each point. When the alpha value is set as 4.2354, we acquire 2810 points on the boundary (Figure 3 (b)). Figure 3 (c) shows the reconstructed tree crown mesh model.

The second example is a candlenut tree without leaves shown in Figure 4. The original data of the candlenut tree has 86675 points (Figure 4 (a), and when the alpha value is set as 0. 41399, 4291 points are left on the boundary. The implementation of our algorithm is shown in Table 1, where the last column is the time spent from data input, to Delaunay triangulation, and to the list of all triangular faces on the boundary.

To show the properness of this approach, the original point model of the candlenut tree is combined to its reconstructed crown mesh model. Figure 4 (b) shows this comparison, and Figure 4 (c) shows a close view of Figure 4 (b). It can be seen in

Table 1. Experimental details on two data sets

Tree	Point set	Alpha value	Points on boundary	Time in secs
Maple	114997	4.2354	2810	1814.16
Candlenut	86675	0.41399	4291	2131.03

Figure 4 (b) and Figure 4 (c) that the reconstructed crown mesh model includes the original point model well.

These two examples show that the shape concavity is well reconstructed. The approach is illustrated here on both dense crown and spare (unfoliaged) one.

6 Conclusion

Current 3D acquisition systems lead to model more and more 3D shapes of real life objects. However, nowadays reconstruction approaches classically fail on high complexity objects, such as trees. Even if nice progresses have been noticed on the main branch structure on un-foliaged trees, the overall reconstruction is not satisfactory, especially on small structures and leaves.

We proposed hereby a method to reconstruct in 3D the scanned tree crown, in order to constrain the definition of the branch structures, especially the thinner ones, and contribute to define local geometrical constraints for leaf area reconstruction.

The principle of our approach is based on the use of the alpha-shape on the range point data set, a generalization of the convex hull and subgraph of the Delaunay triangulation. In the Delaunay triangulation process, we choose the triangle candidates on the boundary according to the alpha value, and constrain the surface mesh to stay a manifold. Therefore, our constructed boundary mesh builds in fact the silhouette of the crown. This shape of the tree crown is much more convincing than the convex hull of the tree crown in keeping the major concave features of the crown. This shape can be used to constrain faithfully the reconstruction of branches and foliage.

The proposed approach was successfully implemented and tested on two data sets.

Of course, the reconstructed crown shape mesh is rough, thus fast to render, and thus not strongly concave, so that higher branching structures are not recreated. In future, progress can be achieved by dividing the data into several subsets according to point density, with different alpha values applied to each subset. Concave silhouette surfaces can then be reconstructed independently, and then merged to a more detailed shape.

It is also interesting to note that such crown shapes do find applications in various domains. Such tree crown can contribute to define intermediate LOD plant models, from real plants or simulated ones. It contributes to define low weighted geometrical models. Of course, appropriate color and transparency value computations can increase the appearance while rendering such shapes.

Finally, the proposed technique may be of interest on a wide range of complex object, showing high topological complexity, where simplified representation, based on internal complex structure is useful. Such could be the case of human organs representation build from their internal vessels.

Acknowledgments

The authors would like to show thanks to Dr. Baoquan CHEN and Ms. Wei Ma for providing scanned data of trees. CGAL library was used for Delaunay triangulation [24]. This work is supported in part by National Natural Science Foundation of China with projects No. 60073007, 60672148, 60473110; in part by the National High

Technology Development 863 Program of China under Grant No.2006AA01Z301; by the French National Research Agency within project NATSIM ANR-05-MMSA-45; and in part by the MOST International collaboration project No. 2007DFC10740.

References

[1] Edelsbrunner, H., Mucke, E.P.: Three dimensional alpha shapes. ACM Trans. Graph. 13, 43–72 (1994)

[2] Jarvis, R.A.: Computing the shape hull of points in the plane. In: Proceedings of IEEE Computer Society Conference on Pattern Recognition and Image Processing, pp. 231–241. IEEE, New York (1977)

[3] Edelsbrunner, H., Kirkpatrick, D.G., Seridel, R.: On the shape of a set of points in the plane. IEEE Trans. Inform. Theory 29(4), 551–559 (1983)

[4] Boissonnat, J.D.: Geometry structure for three-dimensional shape representation. ACM Trans. Graph 3, 266–286 (1984)

[5] Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J., Shade, J., Fulk, D.: The Digital Michelangelo Project: 3D Scanning of Large Statues. In: Proceedings of ACM SIGGRAPH 2000, pp. 131–144 (2000)

[6] de Reffye, P., Edelin, C., Françon, J., Jaeger, M., Puech, C.: Plant models faithful to botanical structure and development. In: SIGGRAPH 1988 Proceedings of the 15th annual conference on Computer graphics and interactive techniques, pp. 151–158. ACM Press, New York (1988)

[7] Prusinkiewicz, P., Lindenmayer, A.: The algorithmic beauty of plants. Springer, New York (1990)

[8] Deussen, O., Lintermann, B.: A modeling method and user interface for creating plants. In: Conference on Graphics interface 1997, Toronto, Ont., Canada, pp. 189–197 (1997)

[9] Xu, H., Gossett, N., Chen, B.Q.: Knowledge-based modeling of laser scanned trees. In: SIGGRAPH 2005: ACM SIGGRAPH 2005 Sketches, p. 124. ACM Press, New York (2005)

[10] Quan, L., Tan, P., Zeng, G., Yuan, L., Wang, J.D., Kang, S.B.: Image-based plant modeling. ACM Trans. Graph. 25(3), 599–604 (2006)

[11] Xu, H., Gossett, N., Chen, B.Q.: Knowledge and heuristic-based modeling of laser-scanned trees. ACM Trans. Graph. 26(4), 19 (2007)

[12] Prusinkiewicz, P., James, M., Mech, R.: Synthetic topiary. In: SIGGRAPH 1994: Proceedings of the 21st annual conference on Computer graphics and interactive techniques, pp. 351–358. ACM Press, New York (1994)

[13] Counède, P.H., Kang, M.Z., Mathieu, A., Yan, H.P., Hu, B.G., de Reffye, P.: Structural factorization of plants to compute their functional and architectural growth, Simulation. Transactions of the Society for Modeling and Simulation International 82(7), 427–438 (2006)

[14] Lintermann, B., Deussen, O.: Interactive modeling of plants. IEEE Comput. Graph. Appl. 19(1), 56–65 (1999)

[15] Bloomenthal, J.: Modeling the mighty maple. In: SIGGRAPH 1985: Proceedings of the 12th annual conference on Computer graphics and interactive techniques, pp. 305–311. ACM Press, New York (1985)

[16] Cheng, Z.L., Zhang, X.P., Chen, B.Q.: Simple reconstruction of tree branches from a single range image. J. Comput. Sci. Technol. 22(6), 846–858 (2007)

[17] Pfeifer, N., Gorte, B., Winterhalder, D.: Automatic reconstruction of single trees from terrestrial laser scanner data. In: Proceedings of 20th ISPRS Congress, pp. 114–119 (2004)
[18] Shlyakhter, I., Rozenoer, M., Dorsey, J., Teller, S.: Reconstructing 3d tree models from instrumented photographs. IEEE Comput. Graph. Appl. 21(3), 53–61 (2001)
[19] Teng, C.H., Chen, Y.S., Hsu, W.H.: Constructing a 3D trunk model from two images. Graph. Models 69(1), 33–56 (2007)
[20] Edelsbrunner, H.: Algorithms in Combinatorial Geometry. EATCS Monographs on Theoretical Computer Science, vol. 10. Springer, Heidelberg (1987)
[21] Preparata, F.P., Shamos, M.I.: Computational Geometry: an Introduction. Springer, New York (1985)
[22] Dey, T.K., Sugihara, K., Bajaj, C.L.: Delaunay triangulations in three dimensions with finite precision arithmetic. Comput. Aided Geom. Design 9(6), 457–470 (1992)
[23] Edelsbrunner, H., Shah, N.R.: Incremental toplogical flipping works for regular triangulations. In: Proceedings of 8th Annual ACM Symposium on Computational. Geometry, pp. 43–52 (1992)
[24] CGAL - Computational Geometry Algorithms Library, `http://www.cgal.org/`

A Survey of Modeling and Rendering Trees

Qi-Long Zhang and Ming-Yong Pang*

Department of Educational Technology, Nanjing Normal University, China
Center for Research of EduGame, Nanjing Normal University, China
panion@netease.com

Abstract. As the representation of the vegetation, trees will be an indispensable part of natural scenes. Many studies have focused on the topic about how to simulate realistic trees efficiently. The goal of this paper is therefore to present an overview of methods applied to modeling and rendering of trees in complex natural scenes. These different types of representations and typical methods used in them are classified and analysized. Finally, we will conclude the paper with possible ideas and key points about modeling and rendering of trees for further research.

Keywords: tree, survey, modeling, rendering.

1 Introduction

Reconstruction of natural scenes has always been the main purpose of computer graphics. Especially, many visualizations applications, such as virtual environment and computer game, occur in natural scenes with trees. In recent years, coupled with the rapid development of Computer Science and technology, it becomes probable for us to create extremely complex outdoor scenes routinely. Use of hardware acceleration also propels the development of rendering trees in realtime. However, trees which are quit different from general regular objects because of their particular properties, are often hard to be constructed with a convincing effect.

2 Related Issues in Representations of Trees

Commonly, the tree composed of a trunk, main branches and leaves, is of great complexity. And forests with large number of trees would be more complex. Obviously, diversity, quantity and complexity of plant organs have an important impact on modeling and rendering of trees. Detailed representations usually mean accurate modeling of trees, which requires large number of primitives. So many data of trees model occupy excessive memory space and calculating times. Even with modern rendering algorithm (including use of hardware acceleration), the time used to render the scene with thousands of trees will not be acceptable. Moreover, the majority of objects in such scene often cover only a few, or even a fraction of pixels on the screen, thus

* Corresponding author.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 757–764, 2008.

leading to use expansive anti-aliasing in the rendering. Lots of applications need to display views at interactive frame rate. And simplified representations of trees could be helpful to fast rendering and avoiding excessive computation, though the poor visual equality is not appropriate for close views. In some cases (during zooming operations or during a walk-through in a natural scene), multiscale models or multiresolution whose complexity can be adapted to their visual importance without losing realistic effects are required.

Addressing these problems above, a large number of methods, which are used for efficiently modeling and simplifying representations of trees, have been proposed by researchers in the last decades. In this paper, we mainly address whole representations of trees, do not consider how to represent details of the trees such as leaves and little branches, which is also an important related topic for trees representations. Also, methods used for modeling and rendering of trees would not be distinguished clearly due to close relation between them.

3 Classification

Those methods used for representations of trees can be categorized according to different dimensions as follows:

(1) Off-line or realtime rendering based on rendering time. (2)static representations or dynamic representations based on adaptive complexity. (3)polygon, image, point set and volume based on render primitive.

There is no single criterion to survey all methods used in modeling and rendering of trees. Based on several criteria above, we will review and discuss various typical methods used in modeling and rendering of trees. Coupled with the distance change between observers and objects in many visualizations applications, appearance of trees may change a lot. The farther the distance is, the more details trees lose. In close view, trees appear with a distinct trunk, main branch and leaves. In far view, trees appear only with a trunk and outline, even a pixel point on screen. Thus, our classification is first related to the level of details (LOD) of trees representations. Usually, static representations mean a fixed LOD which is only appropriate for single view (close or far view), while dynamic representations mean multiresolution or adaptable LOD which is often appropriate for a certain range of views. Obviously, they need different methods to build respectively.

3.1 Static Representations

Commonly, static representations adressing a fixed LOD may use various primitives, such as polygon, image, volume and point set, etc.

3.1.1 Polygon Based Methods

Polygonal and especially triangular models have traditionally been the predominant rendering primitive in computer graphics. Recent developments in hardware acceleration have also focused on triangular data; therefore, this rendering primitive has a certain advantage when it comes to realtime rendering. Many polygonal rendering methods for vegetation exploit generic acceleration methods, such as triangle strips, to

speed up rendering. Since the foliage represents the majority of a tree's geometric complexity, groups of leaves or small branches can be approximated as a single, texture mapped polygon [1].

3.1.2 Image Based Methods

As the most common tool for realtime rendering of forests, billboards use one or several polygons crossing with each other to replace complex geometry [2]. Each polygon has texture that just requires a pre-computation step of the tree images and a certain memory. Thanks to their low cost, they are still considered the best choice in many recent industrial simulators. However, it looks unrealistic while observer moved or walked around the tree. Besides, illumination is stored on the image leading to difficulties for dynamic lighting. For the same reason, animation of branches and leaves is not possible.

The authors in [3], [4] propose a method to improve the view quality of billboards by pre-computing a large set of views i.e. using a whole set of images taken from various view angles. But selecting and blending the images gets costly and could not be done in real-time for a whole forest, and the huge amount of image data may not fit in the graphics memory.

Jakulin proposes a mothod using traditional polygon rendering for the trunk and limbs of a tree, and combines it with IBR for the crown. Indeed, the crown foliage is rendered using multiple parallel layers (slices) in the three orthogonal directions [5]. During preprocessing, several sets of these slices are created from various view points. For each slicing, the primitives (i.e. individual leaves) are assigned to the closest slice. Each slice is then rendered to an individual texture. During rendering, the two slicings that are closest to the actual view direction are rendered simultaneously with correct transparency and blending. The goal of this algorithm was to accommodate architectural walk-through and driving simulations. Because the slicings are perpendicular to the ground, viewing trees directly from above or below the tree is not supported.

Qin et al. proposed another method that can fast render photorealistic images of trees under various kinds of daylight [6]. Their approach is to convert a 3D model of the tree into a representation they named quasi-3D tree. A quasi-3D tree is combined of several 2D buffers. One of them stores geometrical and shading information of tree surfaces, i.e. their normal vectors, relative depth, and shadowing of direct sunlight and skylight, which are used to represent the tree perpendicularly to the ground plane as billboard enriched with depth information, Another 2D buffer may store horizontal mask images for casting shadows. The method could resolve the static lighting of billboards, and make the view look more realistic. Nevertheless, the drawback of billboards will be visible in animations when walking around a tree and walking through or over-viewing at a forest.

3.1.3 Volume Based Methods

Photographs are widely used to record and represent objects in space. The idea that reconstructing and rendering a tree from the photograph would be interesting and challenging. Reche et al. make use of geometric properties of objects in photographs, and reconstruct 3D volume of the tree from photographs triumphantly [7]. On each cell of a recursive grid, they estimate an opacity value and a set of textures from the

calibrated pictures. This opacity value is associated to a set of textures extracted from photographs. For the rendering of the whole tree, the algorithm traverse the cells back-to-front, and render the billboard which is textured by the blend of the two textures corresponding to the two closest views for each cell. Nevertheless, the method just works for the tree with sparse leaves, rather than dense forest, because single pixels contain the blended projection of numerous leaves/branches and background. Another limitation is that capturing the top of the trees is difficult, which is important in landscape rendering with flying movement.

3.1.4 Point Set Based Methods

Reeves introduced the method named Particle Systems [8]. Particle Systems build complex pictures from sets of simple, volume-filling primitives. The whole model of a tree is created firstly, then the rendering of the tree begins with the trunk, and generates sub-branches recursively. The process performs randomly without considering details of the tree. Effects are very surprising, although computing times could be several hours long.

3.2 Dynamic Representations

Dynamic representations which address a range of hierarchical techniques often use several data structures with hierarchical relation, such as binary tree, BSP tree or octree.

3.2.1 Geometry Element Based Methods

Remolar et al. adopt a model in which leaves are composed of unconnected polygons. Therefore, the simplification methods used for traditional triangle mesh model will not be available [9], [10]. They propose a new simplification method. The key operations of their algorithm are leave collapse and split: two leaves are transformed into one with similar area, or one leave is split into two. By preprocessing, a multiresolution model will be created. At the highest resolution, each leave is represented by a polygon, At the coarsest resolusion, root nodes would be the polygons required for simplified representations. The data structure is binary tree.

An error function is defined to decide which pair of leaves will be simplified to create a new one. This function takes into account the distance between two leaves and their planarity. Zhang improved Remolar's model by refining error function with additional criteria such as screen-space projection of local components [1], [10]. Their multiresolution model can represent different part of the model with different resolution.

3.2.2 Adaptable Subdivision Based Methods

Tobler et al. in [12] propose two mechanisms (generalized subdivision and mesh-based parametrized L-Systems) to create smooth mesh for branch. Instead of using standard subdivision, which uses the same subdivision rule at each level of the subdivision process, they employ a generalized approach, that allows different subdivision rules at each level in order to converge to a limit surface,. In the mesh-based parametrized L-Systems, each parameterized symbol represent a face of the mesh. Combining both these mechanisms, a wide variety of complex models can be easily generated from very compact representations. However, it is difficult to create the

initial subdivision mesh for arbitrarily complex branching structures. The author in [13] proposes another improved approach for mesh refinement and growth. And RFL recently apply it to modeling leaves growth [14].

3.2.3 Point Set Based Method

As a development of Reeves' model [8], Weber et al. in [15] and Deussen et al. in [16] propose two models for real time rendering. Their methods combine polygons and points/lines representation as follows: branch meshes as lines and leaf polygons as points. To select which part of the trees has to disappear, Weber et al. use an automatic criterion on the size whereas Deussen et al. ask the user to make the selection during the modeling process.

3.2.4 Fractal Based Methods

Lluch propose a method named procedural multi-resolution based on parametric L-systems which can create parametric string to represent visual structure of the tree at different resolution in realtime [17]. This approach allows the generation of various trees in a forest, but the data storage is still an important issue. In order to avoid storage issue, the generation process will be created only when needed. Digiacomo et al. in [18] extend the idea to the animation and the interaction with trees: procedural method handles most of the trees efficiently, and physically-based method allows user interaction.

IBR can represent complex objects such as trees with a single image, but exist some problems in terms of static illumination, memory cost and reality. An idea that basing several levels of IBR on the hierarchy of tree, is introduced by reasearcher to avoid fixed LOD of images. At coarse level, the whole tree is represented by an IB primitive. At detailed level, each leaves is represented by an IB primitive. The authors in [19], [20], [21], [22] build a simplified tree with a few dozen to a few hundred polygons approximating the foliage distribution. A hierarchy of LOD is built going from a simple billboard to trees with hundred polygons. Behrendt et al. built billboards of a tree by clustering elements of a same level in the hierarchy. In [23] and [4], the IB primitive used in the hierarchy of tree is layered depth image (LDI) and bidirectional texture, respectively.

3.2.5 Space Partition or Sampling Based Methods

Marshall and Fussel in [25] have presented a system for rendering very large collections of randomly parameterized plants. Their multiresolution rendering system compiles plant models into a hierarchical volume approximation based on irregular tetrahedra. This partitioning creates a binary tree similar to BSP trees, which can be traversed quite efficiently. The plant model allows plant information to be stored at various levels of detail and memory usage. The generation of actual geometry for any subvolume can be delayed until it is needed. This drastically reduces memory consumption and initialization time, as the binary tree does not need to be built fully.

The authors in [26], [27] introduce volumetric textures approach which consists of mapping a 3D layer on a surface using a 3D data set as texture pattern. Meyer et al. develop it with hardware acceleration for realtime rendering [28]. In order to represent and render high quality dense forests in real-time, Decaudin et al. in [24]

combine two slicing methods to render volumetric textures efficiently: a simple one used for most locations, and a more complex one used at silhouettes.

According to the complexity and relative size of objects, the space of object is divided into several cubic units [27], [28], [7]. Model in the same cubic unit is replaced with special primitive. The advantage of this method is converting any model input without considering topolopy information. Nevertheless, the model have a poor visual quality in view and difficulty in error control. As preservation of topological relationship can not be guaranteed.

4 Conclusion

Various methods are used to represent realistic trees. Static representations are common addressing several trees representations with fixed level of complexity. Polygon based methods propose interesting techniques, but rely on geometry modification which can produce unadapted results at highest simplification rates. Point based methods offer interesting support to achieve realtime rendering but require hardware support. IBR methods which are usually used for realtime rendering seem to be lack of reality. With the purpose of reducing complexity of natural scenes with trees, Dynamic representations try to improve static representations by defining adaptive model. However, the definition of such progressive simplification scheme is complex on sparse model like trees.

In sum, none of these techniques perfectly answers all. Nevertheless, interesting compromise between realism and efficiency has been proposed and rendering of large landscape is possible. In fact, some points such as animation or trees diversity stay an issue. Moreover, large number of data lead to memory issue. And the massive use of instantiation result in a lack of diversity. As a result, the idea of on the fly data generation begin to be exploited. Point based methods that take full advantage of hardware acceleration seem to be a promiseful method.

Acknowledgments. This work was partly supported by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (Grant No. 07KJD460108), the Outstanding High-end Talent Foundation of Nanjing Normal University (Grant No. 2007013XGQ0150).

References

1. Interactive Data Visualization, Inc. Speedtree product homepage. Web page (2002), `http://www.idvinc.com/speedtree/`
2. Siggraph Course Notes CD-ROM. Advanced Graphics Programming Techniques Using OpenGL. Addison-Wesley (1998)
3. Pulli, K., Cohen, M., Duchamp, T., Hoppe, H., Shapiro, L., Stuetzle, W.: Viewbased rendering: Visualizing real objects from scanned range and color data. In: Eurographics Rendering Workshop (1997)
4. Meyer, A., Neyret, F., Poulin, P.: Interactive rendering of trees with shading and shadows. In: Eurographics Workshop on Rendering (2001)

5. Jakulin, A.: Interactive vegetation rendering with slicing and blending. In: de Sousa, A., Torres, J.C. (eds.) Proceedings of Eurographics (2000)
6. Qin, X., Nakamae, E., Tadamura, K., Nagai, Y.: Fast photorealistic rendering of trees in daylight. In: Computer Graphics Forum Proceedings of Eurographics (2003)
7. Reche, A., Martin, I., Drettakis, G.: Volumetric reconstruction and interactive rendering of trees from photographs. In: ACM Transactions on Graphics, SIGGRAPH 2004 Conference Proceedings (2004)
8. Reeves, W.T., Blau, R.: Approximate and probabilistic algorithms for shading and rendering structured particle systems. In: ACM Transactions on Graphics (SIGGRAPH 1985 Conference Proceedings), pp. 313–322 (1985)
9. Remolar, I., Chover, M., Belmonte, O., Ribelles, J., Rebollo, C.: Geometric simplification of foliage. In: Eurographics 2002 Short Presentation Proceedings, Saarbrcken, Germany, pp. 397–404 (2002)
10. Remolar, I., Chover, M., Ribelles, J., Belmonte, O.: Viewdependent multiresolution model for foliage. Journal of WSCG (WSCG 2003 Proceedings) (2003)
11. Zhang, X.P., Blaise, F.: Progressive polygon foliage simplification. In: Hu, B., Jaeger, M. (eds.) Plant Growth Modeling and Applications (Proceedings of PMA 2003), Beijing, China (2003)
12. Tobler, R.F., Maierhofer, S., Wilkie, A.: Mesh-based parametrized l-systems and generalized subdivision for generating complex geometry. International Journal of Shape Modeling (2002)
13. Smith, C., Prusinkiewicz, P., Samavati, F.: Relational specification of subdivision algorithms. In: Pfaltz, J.L., Nagl, M., Böhlen, B. (eds.) AGTIVE 2003. LNCS, vol. 3062, pp. 313–327. Springer, Heidelberg (2004)
14. Runions, A., Fuhrer, M., Lane, B., Federl, P., Rolland-Lagan, A., Prusinkiewicz, P.: Modeling and visualization of leaf venation patterns. In: ACM Transactions on Graphics (SIGGRAPH 2005 Conference Proceedings), vol. 24(3), pp. 702–711 (2005)
15. Weber, J., Penn, J.: Creation and rendering of realistic trees. In: ACM Transactions on Graphics (SIGGRAPH 1995 Conference Proceedings), pp. 119–128 (1995)
16. Deussen, O., Colditz, C., Stamminger, M., Drettakis, G.: Interactive visualization of complex plant ecosystems. In: Proceedings of the IEEE Visualization Conference, vol. 19, IEEE, October RESEARCH REPORT RR-LIRIS-2006-003, (2002)
17. Lluch, J., Camahort, E., Vivo, R.: Procedural multiresolution for plant and tree rendering. In: Proceedings of the 2nd international conference on Computer graphics, virtual Reality, visualisation and interaction in Africa, pp. 31–38. ACM Press, New York (2003)
18. Di Giacomo, T., Capo, S., Faure, F.: An interactive forest. In: Cani, M.-P., Magnenat-Thalmann, N., Thalmann, D. (eds.) Eurographics Workshop on Computer Animation and Simulation, pp. 65–74. Springer, Manchester (2001)
19. Behrendt, S., Colditz, C., Franzke, O., Kopf, J., Deussen, O.: Realistic real-time rendering of landscapes using billboard clouds. In: Eurographics (2005)
20. Colditz, C., Coconu, L., Deussen, O., Hege, H.: Real-time rendering of complex photorealistic landscapes using hybrid levelof-detail approaches. In: Real-time visualization and participation, 6th International Conference for Information Technologies in Landscape Architecture (2005)
21. Interactive Data Visualization. Speedtree product homepage. web page (2002), `http://www.idvinc.com/speedtree/`
22. Software Blueberry 3D, http://www.blueberry3d.com
23. Max, N., Deussen, O., Keating, B.: Hierarchical image-based rendering using texture mapping hardware. In: Eurographics Workshop on Rendering (1999)

24. Decaudin, P., Neyret, F.: Rendering forest scenes in real-time. In: Keller, A., Jensen, H.W. (eds.) Eurographics Symposium on Rendering (2004)
25. Marshall, D., Fussell, D., Campbell III, A.T.: Multiresolution rendering of complex botanical scenes. In: Davis, W.A., Mantei, M., Klassen, R.V. (eds.) Graphics Interface 1997, pp. 97–104. Canadian Human-Computer Communications Society (1997)
26. Kajiya, J.T., Kay, T.L.: Rendering fur with three dimensional textures. ACM Transactions on Graphics (SIGGRAPH 1989 Conference Proceedings) 23(3), 271–280 (1989)
27. Neyret, F.: Modeling, animating, and rendering complex scenes using volumetric textures. IEEE Transactions on Visualization and Computer Graphics 4(1), 55–70 (1998)
28. Meyer, A., Neyret, F.: Interactive volumetric textures. In: Eurographics Workshop on Rendering, pp. 157–168 (1998)

Creating Boundary Curves of Point-Set Models in Interactive Environment

Pei Xiao and Ming-Yong Pang*

Center for Research on EduGame, Nanjing Normal University, China
Department of Educational Technology, Nanjing Normal University, China
sdxiaopei@163.com, panion@netease.com

Abstract. Extracting boundary curve from point-sampled model is one of key operations in interactive point-cloud modelling. In this paper, an interactive algorithm of creating boundary curves automatically for point-sampled models is proposed based on our analysis about process of selecting boundary points from point-sets in interactive environment. According to a candidate boundary-point set, in which points are selected by user gradually and interactively, our algorithm first calculate the distance from each new candidate point in the set to current boundary, and then add the point to suitable position in the boundary strategically, while the algorithm keeps the integrity of boundary structure and proximately optimizes the length of the boundary curve. The experiment results show that, our method can extract characteristic areas of point-set models or perform model segmentation with general interactive interface efficiently. The algorithm can be used in interactive parameterization and model editing of point-sampled models, etc.

1 Introduction

In past decade, with wide use of 3D data acquisition devices, lots of point-set models [1] arises in application fields, such as digital geometry processing, reverse engineering and virtual reality, etc. At the same time, a variety of techniques for processing point-set models have been proposed by computer graphics scientists, e.g., the algorithms for triangulation of point clouds [2] and surface reconstruction from point-sampled models [3] [4], etc. Usually, the scattered points of a point-set model were sampled from surface of certain physical model by 3D scanner. Thus, the point-set model is a discrete representation of the continuous solid surface. With the enhance of computer processing ability, especially for the use of high-performance graphics acceleration hardware (e.g., GPU), performing various operations on the scatted point-set models directly rather than triangulating them into meshes or piecewise linear surface firstly is becoming more and more popular [2] [5]. These operations include texture-mapping, shape-editing, animation controlling, and so on. As the base of the operations, feature extraction, mesh-less parameterization and segmenting point-set models [6] have been becoming important topics of digital geometry processing.

* Corresponding author.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 765–772, 2008.

At present, researchers have already presented several automatic feature extraction methods and boundary identification algorithms for point-sampled models [7]. These algorithms have ability of analyzing local geometric features, capturing surface boundary and other geometric characters of the models automatically, or segmenting the model surfaces according to user's specific requirements. Although these algorithms have provided basic techniques for automated or heuristic processing of the models, the interactive methods can provide higher degree of freedom for model editing, for example, to interactively outline the model's feature areas or to appropriately change the results created by automatic algorithms. Obviously, the interactive processing technology provides user stronger tools to infiltrate more their personalized modelling intents to editing of geometry and topology of point-set models.

In this paper, we present an algorithm to support generating boundary curves of 3D point-sampled models interactively. Our algorithm can generate optimized boundary curves automatically in the light of a progressive process of selecting boundary points by user. The algorithm can also be further used to cooperate with existing automatic boundary generating techniques to support parameterizing point-clouds, outlining features, segmenting point-set surfaces, etc. The rest of the paper is organized as follows: Section 2 gives the basic idea behind our methods and some detail steps when boundary curve is constructed. In Section 3, our method is implemented and some results are illustrated followed by Section 4, brief conclusion of our work in this paper.

2 Algorithm

The goal of this paper is to offer user an interactive operating mechanism, such that user can deeply participate in extracting boundary curve of a point-sampled model. They can use interactive devices, typically, mouse or keyboard to select the boundary vertices from models interactively, and generate and manage the boundary automatically. Obviously, the boundary can be taken as a kerf, which divides a genus-0 model into two divided parts.

2.1 Data Structure and Its Basic Operation

A boundary curve of point-set model is defined as a closed not-self-intersected polyline, which consists of a series of space points linked by line segments end to end. In order to manage the constructing process of boundary curve, we need appropriate data structure to store the boundary curve information. Since the boundary curve can be uniquely decided by its vertices, only information of the boundary needed to store is its serial vertices, on one hand. In the process of creating boundary, it is frequent occurred thing that a new vertex is added to current boundary or a vertex is removed from the boundary, so dynamic data structure is appropriate for storing the boundary data. In this paper, the double linked list is selected as basic data structure to express the boundary curve loop,

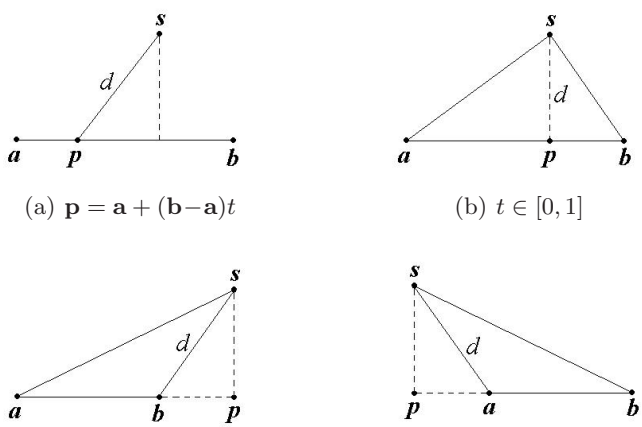
where each node in the list represents one boundary vertex. For the convenience of discussion, we simply name the list data structure above *boundary* or *boundary loop* in this paper.

In interactive environment, with using mouse or other devices, user can pick up some points from a point-sampled model to define a boundary. When gradually adding new point/vertex to the current boundary, one can increase more details of the boundary. If a vertex has been in the boundary, clicking it again by mouse may means removing the vertex from the boundary. Thus, the boundary creating progress involves two basic operations: a) add new vertex to the current boundary; and b) remove a selected vertex from the boundary.

2.2 Adding Vertices to Boundary

Before adding a new vertex to the boundary, the algorithm needs to perform an important step to determine appropriate position in current boundary loop, where the new vertex will be inserted to. This task can be solved by calculating the *distance* between the new vertex and polyline of the current boundary. In this paper, the distance is defined as the shortest one of the distances from the new vertex to all line segments in the boundary curves (the line segments will be called *boundary line segments* or BLS in next sections). Our algorithm will minimize the length of the boundary polyline as optimizing strategy of creating boundary.

As a basic calculating step of our algorithm, we now discuss how to calculate the distance from an arbitrary space point to a BLS of the boundary. Here, let us suppose **a** and **b** are the two adjacent vertices in the current boundary and **s** is the space point in 3D space (see Fig. 1).



(a) $\mathbf{p} = \mathbf{a} + (\mathbf{b} - \mathbf{a})t$

(b) $t \in [0, 1]$

(c) $t > 1$

(d) $t < 0$

Fig. 1. Calculating distance from point **s** to line segment $\overline{\mathbf{ab}}$

Calculate distance between new vertex and boundary. Illustrated as Fig. 1(a), let vector $\mathbf{v} = \mathbf{b} - \mathbf{a}$, obviously, an arbitrary point $\mathbf{p}$ on line L, which is defined by $\overline{\mathbf{ab}}$, can be expressed by a function with respect to t:

$$\mathbf{p} = \mathbf{a} + \mathbf{v}t \tag{1}$$

According to (1), the distance between $\mathbf{p}$ and $\mathbf{s}$ can further be described such as

$$D(t) = ||\mathbf{p} - \mathbf{s}|| = ||\mathbf{a} + \mathbf{v}t - \mathbf{s}|| = ||(\mathbf{a} - \mathbf{s}) + \mathbf{v}t|| \tag{2}$$

Squaring two sides of (2), then expanding and rearrange it, we have:

$$D^2(t) = at^2 + bt + c \tag{3}$$

where, $a = \mathbf{v}^2, b = 2(\mathbf{a} - \mathbf{s}) \cdot \mathbf{v}, c = (\mathbf{a} - \mathbf{s})^2$.

Then, calculating derivation of two sides of Equ.(3),

$$2D'(t) = 2at + b$$

Let $D'(t) = 0$, we can obtain

$$t = \frac{-b}{2a} = \frac{(\mathbf{a} - \mathbf{s}) \cdot (\mathbf{b} - \mathbf{a})}{(\mathbf{b} - \mathbf{a})^2} \tag{4}$$

In this case, the function $D(t)$ get its minimum. Substitute (4) to (2), the distance between $\mathbf{s}$ and L can be evaluated on the fly. However, $D(t)$ must not be the distance between $\mathbf{s}$ and $\overline{\mathbf{ab}}$ in deed. Three special cases must be further analyzed as follows (see Fig.1(b) to 1(d)):

Case 1. In Equ.(4), if $t \in [0, 1]$, point $\mathbf{p}$ is on the segment $\overline{\mathbf{ab}}$ according to the Equ.(1). That is to say, the projection of $\mathbf{s}$ on line L is exactly on $\overline{\mathbf{ab}}$. The distance between $\mathbf{s}$ and $\overline{\mathbf{ab}}$ is $D(t)$ in this case. See Fig.1(b).

Case 2. If $t > 1$, the projection $\mathbf{p}$ is on extension of the line segment $\overline{\mathbf{ab}}$ on the right hand side. Thus, the distance between $\mathbf{s}$ and $\overline{\mathbf{ab}}$ is $d = ||\mathbf{s} - \mathbf{b}||$ rather than $D(t)$ (Fig.1(c)).

Case 3. If $t < 1$, similar to Case 2, we can get the distance between $\mathbf{s}$ and $\overline{\mathbf{ab}}$ as $d = ||\mathbf{s} - \mathbf{a}||$.

Insert new vertex to boundary loop. Once the distances between $\mathbf{s}$ and BLS's in current boundary loop are calculated, we can find the shortest one. If the distance exactly corresponds to one BLS, we denote the BLS by $\overline{\mathbf{v}_i\mathbf{v}_{i+1}}$, the work of adding the new vertex becomes to simply insert the vertex between $\mathbf{v}_i$ and $\mathbf{v}_{i+1}$. Otherwise, there must be several BLS's corresponding to the shortest distance, see Fig. 2(b) for a reference. In this case, inserting the new vertex to the boundary loop is needed to be farther treated carefully:

As showed in Fig. 2(a), assume $\mathbf{v}_{i-1}$, $\mathbf{v}_i$ and $\mathbf{v}_{i+1}$ are neighboring vertices in the current boundary loop. There are two planes passing $\mathbf{v}_i$ and perpendicular to $\overline{\mathbf{v}_{i-1}\mathbf{v}_i}$ and $\overline{\mathbf{v}_i\mathbf{v}_{i+1}}$ respectively, which bound a space domain illustrated by the

shadowed area in the figure. If $\mathbf{s}$ is in the domain C, the distance between $\mathbf{s}$ and $\overline{\mathbf{v}_{i-1}\mathbf{v}_i}$ is equal to the distance between $\mathbf{s}$ and $\overline{\mathbf{v}_i\mathbf{v}_{i+1}}$. In this situation, according to Equ.(4), we know $t_{i-1} > 1$ for $\overline{\mathbf{v}_{i-1}\mathbf{v}_i}$ and $t_i < 0$ for $\overline{\mathbf{v}_i\mathbf{v}_{i+1}}$. This result can inversely be used to decide if $\mathbf{s}$ is in the shadowed area or not. Moreover, we denote by π the plane, which equably bisect the angle $\angle\mathbf{v}_{i-1}\mathbf{v}_i\mathbf{v}_{i+1}$ (see Fig.2(a)). If $\mathbf{s}$ and $\mathbf{v}_{i-1}$ are in the same side of π, insert $\mathbf{s}$ into $\overline{\mathbf{v}_{i-1}\mathbf{v}_i}$, otherwise insert $\mathbf{s}$ into $\overline{\mathbf{v}_i\mathbf{v}_{i+1}}$. This approach can avoid serious distortion of boundary curve in local area. If $\mathbf{s}$ is exactly in the plane π, it is needed first to compare the distances from $\mathbf{s}$ to $\mathbf{v}_{i-1}$ and $\mathbf{v}_{i+1}$, if $||\mathbf{s}-\mathbf{v}_{i-1}|| \leq ||\mathbf{s}-\mathbf{v}_{i+1}||$, insert $\mathbf{s}$ between $\mathbf{v}_{i-1}$ and $\mathbf{v}_i$, otherwise, between $\mathbf{v}_i$ and $\mathbf{v}_{i+1}$. By this way, we can make the boundary curve as short as possible.

In the process above, evaluating the equation of plane π is needed. Since $\mathbf{v}_i$ is a point of π, the only thing for evaluating the equation is to decide the normal $\mathbf{n}$ of π. In the plane defined by $\mathbf{v}_{i-1}\mathbf{v}_i\mathbf{v}_{i+1}$, the direction of the bisector of $\triangle\mathbf{v}_{i-1}\mathbf{v}_i\mathbf{v}_{i+1}$ can be expressed as:

$$\mathbf{m}_1 = \frac{\mathbf{v}_{i-1}-\mathbf{v}_i}{||\mathbf{v}_{i-1}-\mathbf{v}_i||} + \frac{\mathbf{v}_{i+1}-\mathbf{v}_i}{||\mathbf{v}_{i+1}-\mathbf{v}_i||}$$

On the other hand, the normal of plane $\mathbf{v}_{i-1}\mathbf{v}_i\mathbf{v}_{i+1}$ is:

$$\mathbf{m}_2 = \frac{\mathbf{v}_{i-1}-\mathbf{v}_i}{||\mathbf{v}_{i-1}-\mathbf{v}_i||} \times \frac{\mathbf{v}_{i+1}-\mathbf{v}_i}{||\mathbf{v}_{i+1}-\mathbf{v}_i||}$$

Obviously, $\mathbf{n}//(\mathbf{m}_1\times\mathbf{m}_2)$, that is to say, the equation of plane π can be written as:

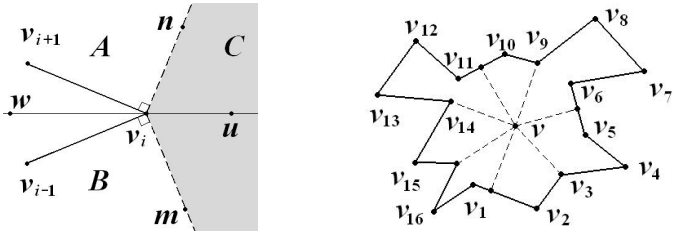
$$\left\{\left(\frac{\mathbf{v}_{i-1}-\mathbf{v}_i}{||\mathbf{v}_{i-1}-\mathbf{v}_i||}+\frac{\mathbf{v}_{i+1}-\mathbf{v}_i}{||\mathbf{v}_{i+1}-\mathbf{v}_i||}\right)\times\left(\frac{\mathbf{v}_{i-1}-\mathbf{v}_i}{||\mathbf{v}_{i-1}-\mathbf{v}_i||}\times\frac{\mathbf{v}_{i+1}-\mathbf{v}_i}{||\mathbf{v}_{i+1}-\mathbf{v}_i||}\right)\right\}\cdot(\mathbf{v}-\mathbf{v}_i)=0$$

An extreme case is illustrated in Fig.2(b), in which the shortest distance corresponds to a set of BLS's. Here, let Ω is the set consisted of the BLS's. The first step we do is to find the BLS's that satisfy $t \in [0,1]$ in Equ.(4) from Ω. For example, $\overline{\mathbf{v}_1\mathbf{v}_2}$ and $\overline{\mathbf{v}_5\mathbf{v}_6}$ in the figure. And then, in these BLS's, to search the BLS, which satisfies that the distance sum from $\mathbf{s}$ ($\mathbf{v}$ in the figure) to its two ends is minimum. Without loss of generality, we denote the found BLS by $\overline{\mathbf{v}_k\mathbf{v}_{k+1}}$, we now can insert the new vertex between $\mathbf{v}_k$ and $\mathbf{v}_{k+1}$ to keep the total length of the boundary optimized.

If there is no BLS satisfied $t \in [0,1]$ in Ω, we can assert that each BLS in Ω belongs to the pattern illustrated in Fig.2(a), i.e., each BLS has one end apart from $\mathbf{s}$ the shortest distance. In this situation, evaluate distances from $\mathbf{s}$ to the other end of each BLS, then take the shortest one as the candidate BLS and insert the new vertex between its ends.

2.3 Removing Vertex from Boundary

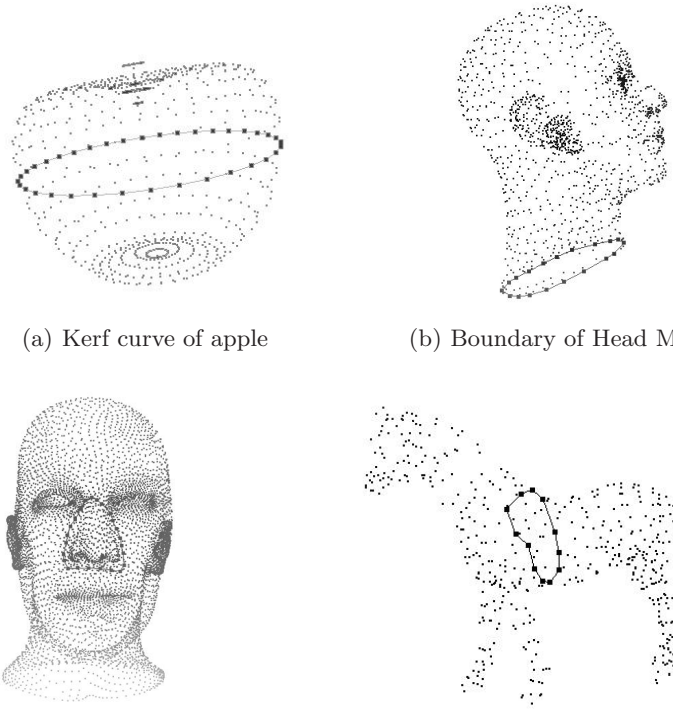
During the process of creating boundary curve, once we found that one vertex in current boundary was not selected appropriately, with clicking it again by



(a) an local amphibolous situation

(b) an extreme example

Fig. 2. Dealing with special cases



(a) Kerf curve of apple

(b) Boundary of Head Model

(c) Outline nose feature

(d) Kerf cure on horse

Fig. 3. Some results of our algorithm

mouse, we can delete it from the boundary. In this case, we only need to remove the corresponding node from the list.

3 Implementation and Results

We implement the algorithm by C++ on PC. In order to make the algorithm can be transplanted on cross-platform, we employ the object-oriented programming

skills. The data structure and related interactive operation are encapsulated in separate classes, so the functions of our implementation is independence to particular window system. On different GUI platform, one can manage the point-set boundary curve through calling the interface functions of the classes of the algorithm.

In our experiments, we take GLUT library as window system to visualize the experimental results. Through the link between the algorithm and the callbacks of display window, keyboard and mouse, our tester can carry selecting and deleting of the boundary vertex. At the same time, we use OpenGL to accomplish 3D transformation of point-set models [8]. Picking and selecting functions are implemented by OpenGL API.

Fig.3 shows some results of creating boundary curves interactively based on our algorithm. Fig.3(a) shows the boundary curve generating result of the apple model which is composed by 867 sample points, the boundary divided the apple model into two parts. Fig.3(b) use the algorithm creating manhead model section curve, this curve can be used as the boundary of plane parameter domain during point-set model parameterizing. Fig.3(c) selects the nose region of the manhead, the related boundary curve can be used in local editing of the model. Fig.3(d) shows the situation of cross-section of the simplified horse point-set model, the relevant curve is used to calculating the parameterization on sphere surface of the model.

4 Conclusion

In this paper, an algorithm that can interactively generates boundary curves for point-set models was proposed. According to a candidate boundary-point set, in which points are selected by user gradually, calculate the distance from each new candidate point in the set to current boundary and proximately optimizes the length of the boundary curve. Cooperate with OpenGL and windows system, we implement the algorithm, and apply it successfully in parameterization, feature designate, point-set segmentation.

Acknowledgement

This work was partly supported by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (Grant No. 07KJD460108), the Outstanding High-end Talent Foundation of Nanjing Normal University (Grant No. 2007013XGQ0150).

References

1. Egenhofer, M.J., Franzosa, R.D.: Point-Set Topological Spatial Relations. International Journal for Geographical Information Systems 5(2), 161–174 (1991)
2. Hormann, K., Labsik, U., Greiner, G.: Remeshing triangulated surfaces with optimal parameterizations. Computer-Aided Design 33(11), 779–788 (2001)

3. Floater, M.S., Hormann, K., Reimers, M.: Parameterization of manifold triangulations. In: Chui, C.K., Schumaker, L.L., Stockler, J. (eds.) Approximation Theory X: Abstract and Classical Analysis, Innovations in Applied Mathematics, pp. 197–209. Vanderbilt University Press, Nashville (2002)
4. Hormann, K.: From scattered samples to smooth surfaces. In: Proc. of the 4th Israel-Korea Bi-National Conf. on Geometric Modeling and Computer Graphics, pp. 1–5 (2003)
5. Xiao, C.X., Zheng, W.T., Peng, Q.S.: Robust morphing of point-sampled geometry. Computer Animation and Virtual Worlds, 201–210 (Special Issue, 2004)
6. Woo, H., Kang, E.: A new segmentation method for point cloud data. International Journal of Machine Tools & Manufacture 42, 167–178 (2002)
7. Shan, D.R., Ke, Y.L.: Quadric Feature Extraction from oints Cloud in Reverse Engineering (in Chinese). Journal of Computer Aided Design & Computer Graphics 15(12), 1497–1501 (2003)
8. Pang, M.Y., Lu, Z.P.: Implementation of Simple Class of Mouse Trackball (in Chinese). Computer Engineering 30(17), 82–183 (2004)

Rational Biquartic Interpolating Surface Based on Function Values

Siqing Deng[1], Kui Fang[2,3], Jin Xie[4], and Fulai Chen[1]

[1] Dep. of Math., Xiangnan Univ., Chenzhou 423000, Hunan, China
[2] Sch. of Info. Sci. & Tech., Hunan Agricultural Univ., Changsha 410128, Hunan, China
[3] Sch. of Math. & Computer, Hunan Normal Univ., Changsha 410081, Hunan, China
[4] Dep. of Math. & Physics, Hefei Univ., Hefei 230601, Anhui, China
dengsq66@163.com

Abstract. In this paper a bivariate rational biquartic interpolating spline based on function values with two parameters is constructed , and this spline is with biquartic numerator and bilinear denominator. The interpolating function has a simple and explicit mathematical representation, which is convenient both in practical application and in theoretical study. The interpolating surface is C^1 in the interpolating region when one of the parameters satisfies a simple condition. The interpolating surface can be modified by selecting suitable parameters under the condition that the interpolating data are not changed. It is proved that the values of the interpolating function in the interpolating region are bounded no matter what the parameters might be; this is called the bounded property of the interpolation. The approximation expressions of the interpolation are derived:they do not depend on the parameters.

Keywords: Bivariate interpolation; Bivariate spline; Rational spline; Parameter; Computer aided geometric design.

1 Introduction

The construction method of the curve and surface and the mathematical description of them is a key issue in computer aided geometric design. There are many ways to tackle this problem [1-13], for example, the polynomial spline method, the Non-Uniform Rational B-Spline(NURBS)method, and the Bézier method. These methods are effective and applied widely in shape design of industrial products, such as aircraft and ships. Generally speaking, most of the polynomial spline methods are interpolating methods, which means that the curves or surfaces constructed by those methods pass through the interpolant points. To construct the polynomial spline, the derivative values are usually needed , as well as the function values as interpolating data. Unfortunately, in many practical problems, such as the description of the rainfall in some rainy region and some geometric shapes, the derivative values are difficult to get. On the other hand, one of the disadvantages of the polynomial spline method is its global property; it is not possible for local modification under the condition that the given data are not changed.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 773–780, 2008.

The NURBS and Bézier methods are the so-called "no-interpolating type" methods. This means that the constructed curve and surface do not pass through the given data, and the given points play the role of the control points. Thus, constructing the interpolating function, which satisfies the following conditions, will be necessary in CAGD: there are only function values as the interpolating data; the interpolating functions have simple and explicit representations, so they may be convenient to use both in practical applications and in theoretical study; the constructed curves and surfaces can be modified under the conditions that the given data are not changed.

In recent years, univariate rational spline interpolations with parameters have been constructed[14-15]. Motivated by the univariate rational spline interpolation, the bivariate rational bicubic interpolation with parameters, based only on the values of the function being interpolated, and some properties of the interpolation have been studied in [16]. The bivariate rational bicubic interpolation based on function values and partial derivatives with parameters have been constructed, and some properties of the interpolation have been studied in [17]. A bivariate rational biquartic interpolation have been constructed [20]. This paper will deal with the new bivariate rational biquartic interpolation based on function values of the function being interpolated.

The paper is arranged as follows. In Section 2, the new bivariate rational biquartic spline based on function values with two parameters is constructed , and this spline is with biquartic numerator and bilinear denominator. Section 3 deals with the smoothness of the interpolating surfaces, when one of the parameters satisfies a simple condition: the interpolating function is C^1 in the interpolating region. In Section 4, the basis of the interpolation are derived. Section 5 deals with bounded property and approximation of the interpolation. For the given interpolation data, it is proved that the values of the interpolating function in the interpolating region are bounded no matter what the parameters might be; this is called the bounded property of the interpolation. Also, the approximation expressions of the interpolation are derived.

2 Interpolation

Let $\Omega:[a,b;c,d]$ be the plane region, and $\{(x_i, y_j, f_{i,j}), i=1,2,\cdots,n,n+1;$ $j=1,2,\cdots,m,m+1\}$ be a given set of data points, where $a=x_1<x_2<\cdots<x_n<x_{n+1}=b$ and $c=y_1<y_2<\cdots<y_m<y_{m+1}=d$ are the knot spacings. Let $h_i=x_{i+1}-x_i$, and $l_j=y_{j+1}-y_j$, for any point $(x,y)\in[x_i,x_{i+1};y_j,y_{j+1}]$ in the $xy-$ plane, and let $\theta=(x-x_i)/h_i$ and $\eta=(y-y_j)/l_j$. First, for each $y=y_j, j=1,2,\cdots,m+1$, construct the $x-$direct interpolant curve as follows:

$$P^*_{i,j}(x)=\frac{p^*_{i,j}(x)}{q^*_{i,j}(x)},\quad i=1,2,\cdots,n-1, \tag{1}$$

where

$$p^*_{i,j}(x)=\alpha_{i,j}f_{i,j}(1-\theta)^4+U^*_{i,j}\theta(1-\theta)^3+V^*_{i,j}\theta^2(1-\theta)^2+W^*_{i,j}\theta^3(1-\theta)+f_{i+1,j}\theta^4,$$
$$q^*_{i,j}(x)=\alpha_{i,j}(1-\theta)+\theta,$$

and

$$U_{i,j}^{*}=(2\alpha_{i,j}+1)f_{i,j}+\alpha_{i,j}f_{i+1,j},\ V_{i,j}^{*}=3f_{i,j}+3\alpha_{i,j}f_{i+1,j},$$
$$W_{i,j}^{*}=(\alpha_{i,j}+3)f_{i+1,j}-h_i\Delta_{i+1,j}^{*}.$$

with $\alpha_{i,j}>0$, and $\Delta_{i,j}^{*}=(f_{i+1,j}-f_{i,j})/h_i$. This interpolation is called the rational quartic interpolation based on function values which satisfies

$$P_{i,j}^{*}(x_i)=f_{i,j},\ \ P_{i,j}^{*}(x_{i+1})=f_{i+1,j},\ \ \frac{dP_{i,j}^{*}(x)}{dx}\Big|_{x=x_i}=\Delta_{i,j}^{*},\ \frac{dP_{i,j}^{*}(x)}{dx}\Big|_{x=x_{i+1}}=\Delta_{i+1,j}^{*}.$$

Obviously, the interpolating function $P_{i,j}^{*}(x)$ on $[x_i,x_{i+1}]$ is existent for the given data $\{x_r,f(x_r,y_j)\},r=i,i+1,i+2$ and parameter $\alpha_{i,j}$.

For each pair of (i,j), $i=1,2,\cdots,n-1$ and $j=1,2,\cdots,m-1$, using the $x-$direct interpolation function $P_{i,j}^{*}(x)$, define the bivariate rational biquartic interpolating function $P_{i,j}(x,y)$ on $[x_i,x_{i+1};y_j,y_{j+1}]$ as follows:

$$P_{i,j}(x,y)=\frac{p_{i,j}(x,y)}{q_{i,j}(y)},\ i=1,2,\cdots,n-1\ ;\ j=1,2,\cdots,m-1, \tag{2}$$

where

$$p_{i,j}(x,y)=\beta_{i,j}P_{i,j}^{*}(x)(1-\eta)^4+U_{i,j}\eta(1-\eta)^3+$$
$$V_{i,j}\eta^2(1-\eta)^2+W_{i,j}\eta^3(1-\eta)+P_{i,j+1}^{*}(x)\eta^4,$$
$$q_{i,j}(y)=\beta_{i,j}(1-\eta)+\eta,$$

and

$$U_{i,j}=(2\beta_{i,j}+1)P_{i,j}^{*}(x)+\beta_{i,j}P_{i,j+1}^{*}(x),\quad V_{i,j}=3P_{i,j}^{*}(x)+3\beta_{i,j}P_{i,j+1}^{*}(x),$$
$$W_{i,j}=(\beta_{i,j}+3)P_{i,j+1}^{*}(x)-l_j\Delta_{i,j+1}(x),$$

with $\beta_{i,j}>0$, and $\Delta_{i,j}(x)=(P_{i,j+1}^{*}(x)-P_{i,j}^{*}(x))/l_j$. Therefore, $P_{i,j}(x,y)$ is called the bivariate rational biquartic interpolating function based on function values which satisfies

$$P_{i,j}(x_r,y_s)=f(x_r,y_s),\quad r=i,i+1,\quad s=j,j+1.$$

It is easy to understand that interpolating function $P_{i,j}(x,y)$ on $[x_i,x_{i+1};y_j,y_{j+1}]$ is existent for the given data $(x_r,y_s,f(x_r,y_s))$, $r=i,i+1,i+2$, $s=j,j+1,j+2$ and parameters $\alpha_{i,j},\beta_{i,j}$.

3 Condition for C^1 Interpolatory Surface

The rational interpolant function $P_{i,j}^{*}(x)$ defined by (1) has continuous first-order derivative when $x\in[x_1,x_n]$, so ti is easy to see that the bivariable interpolant function

$P_{i,j}(x,y)$ defined by (2) has continuous first-order partial derivative $\frac{\partial P_{i,j}(x,y)}{\partial y}$ and $\frac{\partial P_{i,j}(x,y)}{\partial x}$ in the interpolating region $[x_1,x_n;y_{1,}y_m]$ except for every $y\in[y_j,y_{j+1}]$, $j=1,2,\cdots,m-1$ at the points (x_i,y), $i=2,3,\cdots,n-1$, so it is sufficient for $P_{i,j}(x,y)\in C^1$ in the whole interpolating region $[x_1,x_{n-1};y_{1,}y_{m-1}]$ is to find the condition that $\frac{\partial P_{i,j}(x_i+,y)}{\partial x}=\frac{\partial P_{i,j}(x_i-,y)}{\partial x}$ holds. This leads to the following theorem.

Theorem 1. The sufficient condition for the interpolating function $P_{i,j}(x,y)$, $i=1,2,\cdots,n$; $j=1,2,\cdots,m$ to be C^1 in the whole interpolating region $[x_1,x_n;y_{1,}y_m]$ is the parameters $\beta_{i,j}=cons\tan t$, for each $j\in\{1,2,\cdots,m-1\}$ and all $i=1,2,\cdots,n-1$.

Proof. Without loss of generality, for any pare real numbers of $(i,j),1\le i\le n-1$, $1\le j\le m-1$ and $y\in[y_j,y_{j+1}]$, it is sufficient to prove that $\frac{\partial P_{i,j}(x_i+,y)}{\partial x}=\frac{\partial P_{i,j}(x_i-,y)}{\partial x}$.

Since

$$\frac{\partial P_{i,j}(x,y)}{\partial x}=\frac{1}{q_{i,j}(y)}[\beta_{i,j}\frac{dP^*_{i,j}(x)}{dx}(1-\eta)^4+\frac{dU_{i,j}}{dx}\eta(1-\eta)^3+\frac{dV_{i,j}}{dx}\eta^2(1-\eta)^2$$

$$+\frac{dW_{i,j}}{dx}\eta^3(1-\eta)+\frac{dP^*_{i,j+1}(x)}{dx}\eta^4],$$

and

$$P^{*\prime}_{i,r}(x_i+)=\Delta^*_{i,r},\ r=j,j+1,j+2,$$

so

$$U'_{i,j}(x_i+)=(2\beta_{i,j}+1)\Delta^*_{i,j}+\beta_{i,j}\Delta^*_{i,j+1},\ V'_{i,j}(x_i+)=3\Delta^*_{i,j}+3\beta_{i,j}\Delta^*_{i,j+1},$$

$$W'_{i,j}(x_i+)=(\beta_{i,j}+3)\Delta^*_{i,j+1}-\frac{l_j}{l_{j+1}}(\Delta^*_{i,j+2}-\Delta^*_{i,j+1})\cdot$$

and thus

$$\left.\frac{\partial P_{i,j}(x,y)}{\partial x}\right|_{x=x_i+}=\frac{Q_1(\eta)}{\beta_{i,j}(1-\eta)+\eta},\tag{3}$$

where

$$Q_1(\eta)=\beta_{i,j}\Delta^*_{i,j}(1-\eta)^4+((2\beta_{i,j}+1)\Delta^*_{i,j}+\beta_{i,j}\Delta^*_{i,j+1})\eta(1-\eta)^3+(3\Delta^*_{i,j}+3\beta_{i,j}\Delta^*_{i,j+1})\eta^2(1-\eta)^2$$
$$+((\beta_{i,j}+3)\Delta^*_{i,j+1}-\frac{l_j}{l_{j+1}}(\Delta^*_{i,j+2}-\Delta^*_{i,j+1}))\eta^3(1-\eta)+\Delta^*_{i,j+1}\eta^4\cdot$$

Similarly, since $P^{*\prime}_{i-1,r}(x_i-)=\Delta^*_{i,r}$, $r=j,j+1,j+2$, it can be shown that

$$\left.\frac{\partial P_{i-1,j}(x,y)}{\partial x}\right|_{x=x_i-}=\frac{Q_2(\eta)}{\beta_{i-1,j}(1-\eta)+\eta}, \tag{4}$$

where

$$Q_2(\eta)=\beta_{i-1,j}\Delta^*_{i,j}(1-\eta)^4+((2\beta_{i-1,j}+1)\Delta^*_{i,j}+\beta_{i-1,j}\Delta^*_{i,j+1})\eta(1-\eta)^3+(3\Delta^*_{i,j}+3\beta_{i-1,j}\Delta^*_{i,j+1})\eta^2(1-\eta)^2$$
$$+((\beta_{i-1,j}+3)\Delta^*_{i,j+1}-\frac{l_j}{l_{j+1}}(\Delta^*_{i,j+2}-\Delta^*_{i,j+1}))\eta^3(1-\eta)+\Delta^*_{i,j+1}\eta^4\cdot$$

Comparing (3) and (4), if $\beta_{i-1,j}=\beta_{i,j}$, then

$$\frac{\partial P_{i,j}(x_i+,y)}{\partial x}=\frac{\partial P_{i,j}(x_i-,y)}{\partial x}.$$

4 The Basis of the Interpolation

In what follows in this paper, consider the equally spaced knots case, namely,for all $i=1,2,\cdots,n$ and $j=1,\ 2,\ \cdots,m$, $h_i=h_j$,denote it by h, and $l_i=l_j$, denote it by l. From Theorem 1, the C^1 -continuous interpolation defined by (1) must satisfy $\beta_{i,j}=cons\tan t$, for each $j\in\{1,2,\cdots,m-1\}$ and all $i=1,2,\cdots,n-1$. Denote this value by β_j. In the following the case that $\alpha_{i,j}=cons\tan t$, for each $i\in\{1,2,\cdots,n-1\}$ and all $j=1,2,\cdots,m-1$ is considered, and denote it by α_i. Under the conditions above, $P^*_{i,j}(x)$ defined by (1) could be rewritten as

$$P^*_{i,j}(x)=\omega_0(\theta,\alpha_i)f_{i,j}+\omega_1(\theta,\alpha_i)f_{i+1,j}+\omega_2(\theta,\alpha_i)f_{i+2,j}, \tag{5}$$

where

$$\omega_0(\theta,\alpha_i)=\frac{\alpha_i(1-\theta)^3(1+\theta)+\theta(1-\theta)^2(1+2\theta)}{\alpha_i(1-\theta)+\theta},$$

$$\omega_1(\theta,\alpha_i)=\frac{\alpha_i\theta(1-\theta)(1+\theta(1-\theta))+4\theta^3-3\theta^4}{\alpha_i(1-\theta)+\theta},\ \omega_2(\theta,\alpha_i)=\frac{-\theta^3(1-\theta)}{\alpha_i(1-\theta)+\theta}\cdot$$

and

$$\sum_{r=0}^{2}\omega_r(\theta,\alpha_i)\equiv 1\cdot$$

Similarly, the bivariate rational interpolating function $P_{i,j}(x,y)$ defined by (2) can be expressed as the following:

$$P_{i,j}(x,y)=\sum_{r=0}^{2}\sum_{s=0}^{2}\omega_{rs}(\theta,\alpha_i;\eta,\beta_j)f_{i+r,j+s} \tag{6}$$

where

$$\omega_{rs}(\theta,\alpha_i;\eta,\beta_j)=\omega_r(\theta,\alpha_i)\omega_s(\eta,\beta_j) \tag{7}$$

and

$$\sum_{r=0}^{2}\sum_{s=0}^{2}\omega_{rs}(\theta,\alpha_i;\eta,\beta_j)\equiv 1\cdot \tag{8}$$

5 Bounded Property and Approximation of the Interpolation

For the given data, the values of the bivariate interpolating function defined by (2) are bounded in the interpolation interval as described by the following Theorem 2.

Theorem 2. Let $P_{i,j}(x,y)$ be the bivariate interpolating function defined by (2) in $[x_i,x_{i+1};y_j,y_{j+1}]$, and denote

$$M=\max_{r=0,s=0}^{r=2,s=2}\left|f_{i+r,j+s}\right|\cdot$$

Whatever the positive values of the parameters α_i and β_j might be, the values of $P_{i,j}(x,y)$ in $[x_i,x_{i+1};y_j,y_{j+1}]$ satisfy

$$\left|P_{i,j}(x,y)\right|\le\frac{1225}{729}M\cdot$$

Proof. From (6) and (7),

$$\left|P_{i,j}(x,y)\right|\le\sum_{r=0}^{2}\sum_{s=0}^{2}\left|\omega_{rs}(\theta,\alpha_i;\eta,\beta_j)\right|\left|f_{i+r,j+s}\right|\le M\sum_{r=0}^{2}\left|\omega_r(\theta,\alpha_i)\right|\sum_{s=0}^{2}\left|\omega_s(\eta,\beta_j)\right|\cdot \tag{9}$$

When $\theta\in[0,1]$, $\omega_0(\theta,\alpha_i)\ge 0$, $\omega_1(\theta,\alpha_i)\ge 0$ and $\omega_2(\theta,\alpha_i)\le 0$, so, it is easy to show that

$$\sum_{r=0}^{2}\left|\omega_r(\theta,\alpha_i)\right|=1+\frac{2\theta^3(1-\theta)}{\alpha_i(1-\theta)+\theta}\cdot \tag{10}$$

Similarly,

$$\sum_{s=0}^{2}\left|\omega_s(\eta,\beta_j)\right|=1+\frac{2\eta^3(1-\eta)}{\beta_j(1-\eta)+\eta}\cdot \tag{11}$$

Denote

$$g(\theta)=1+\frac{2\theta^3(1-\theta)}{\alpha_i(1-\theta)+\theta},$$

since for any $\alpha_i>0$ and $\theta\in[0,1]$, $g(\theta)\le 1+2\theta^2(1-\theta)$, and since $\max_{\theta\in[0,1]}g(\theta)\le\frac{35}{27}$, no matter what the positive parameters α_i and β_j might be,

$$\sum_{r=0}^{2}\left|\omega_r(\theta,\alpha_i)\right| \le \frac{35}{27} \text{ and } \sum_{s=0}^{2}\left|\omega_s(\eta,\beta_j)\right| \le \frac{35}{27} \tag{12}$$

so

$$\left|P_{i,j}(x,y)\right| \le (\frac{35}{27})^2 M = \frac{1225}{729} M \cdot$$

Consider the approximation of the interpolation. Denote

$$\left\|\frac{\partial f}{\partial x}\right\| = \max_{x\in[x_1,x_{n+1}]}\left|\frac{\partial f(x,y)}{\partial x}\right|, \quad \left\|\frac{\partial f}{\partial y}\right\| = \max_{y\in[y_1,y_{m+1}]}\left|\frac{\partial f(x,y)}{\partial y}\right| \cdot$$

Theorem 3. Let $P_{i,j}(x,y)$ be the bivariate interpolating function defined by (2) in $[x_i,x_{i+1};y_j,y_{j+1}]$, Whatever the positive values of the parameters α_i and β_j might be, the error of the interpolation satisfies

$$\left|f(x,y)-P_{i,j}(x,y)\right| \le \frac{2450}{729}(h\left\|\frac{\partial f}{\partial x}\right\| + l\left\|\frac{\partial f}{\partial y}\right\|) \cdot$$

6 Conclusion

For bivariate interpolation in general, finding the bounds of the values of the interpolating function expressed by the interpolation data is hard work, and deriving the error estimate formula of the bivariate interpolation function is much more difficult. In this paper, two of them are worked out in Theorem 2 and Theorem 3, respectively. This is because of the convenient basis of the interpolation. The basis are very useful both in practical design and in theoretical study.

Acknowledgements. The supports of the National Nature Science Foundation of China, the Natural Science Foundation of Hunan Province of China and the Research Project of Department of Education of Hunan Province of China are gratefully acknowledged.

References

1. Farin, G.: Curves and surfaces for computer aided geometric design: A practical guide. Academic Press, London (1988)
2. Chui, C.K.: Multivariate spline. SIAM, Philadelphia (1988)
3. Bézier, P.E.: The mathematical basis of the UNISURF CAD system. Butterworth, London (1986)
4. Dierck, P., Tytgat, B.: Generating the Bézier points of β-spline curve. Computer Aided Geometric Design 6, 279–291 (1989)
5. Piegl, L.: On NURBS: A survey. IEEE Computer Graphics and Application 11, 55–71 (1991)
6. Mielson, N.G.: CAGD's Top Ten: What to watch. IEEE Computer Graphics and Automation, 35–37 (1993)

7. Konno, K., Chiyokura, H.: An approach of designing and controlling free-form surfaces by using NURBS boundary Gregory patches. Computer Aided Geometric Design 13, 825–849 (1996)
8. Wilcox, L.M.: First and second contributions to surface interpolation. Vision Research 39, 2335–2347 (1999)
9. Lin, R.S.: Real-time surface interpolator for 3-D parametric surface machining on 3-axis machine tools. Machine tools and Manufacture 40, 1513–1526 (2000)
10. Jiang, D., Liu, H., Wang, W.: Test a modified surface wind interpolation scheme for complex terrain in a stable atmosphere. Atmospheric Environment 35, 4877–4885 (2001)
11. Comninos, P.: An interpolating piecewise bicubic surface with shape parameters. Computer and Graphics 25, 463–481 (2001)
12. Müller, R.: Universal parametrization and interpolation on cubic surfaces. Computer Aided Geometric Design 19, 479–502 (2002)
13. Wang, R.H.: Multivariats Spline Functions and Their Application. Science press/Kluwer Academic, Beijing (2001)
14. Duan, Q., Djidjeli, K., Price, W.G., et al.: A rational cubic spline based on function values. Computer and Graphics 22, 479–486 (1998)
15. Duan, Q., Djidjeli, K., Price, W.G., et al.: The approximation properties of some rational cubic splines. International J. of Computer Mathematics 72, 155–166 (1999)
16. Duan, Q., Wang, L., Twizell, E.H.: A new bivariate rational interpolation based on function values. Information Sciences 166, 181–191 (2004)
17. Duan, Q., Zhang, Y., Twizell, E.H.: A bivariate rational interpolation and the properties. Applied Mathematics and Computation 179, 190–199 (2006)
18. Duan, Q., Zhang, H., Zhang, Y., et al.: Bounded property and point control of a bivariate rational interpolating surface. Computers and Mathematics with Applications 52, 975–984 (2006)
19. Wang, Q., Tan, J.: Rational quartic spline involving shape parameters. Journal of Information and Computational Science 1, 127–130 (2004)
20. Wang, Q., Tan, J.: Shape preserving piecewise rational biquartic surface. Journal of Information and Computational Science 3, 295–302 (2006)

3D Modelling for Metamorphosis for Animation

Li Bai[1], Yi Song[2], and Yangsheng Wang[3]

[1] School of Computer Science, University of Nottingham, Nottingham, UK
[2] School of Computing, University of Leeds, Leeds, UK
[3] Institute of Automaton, Chinese Academy of Sciences, Beijing, China
bai@cs.nott.ac.uk, yisong@comp.leeds.ac.uk,
yangsheng.wang@ia.ac.cn

Abstract. In this paper, we propose a novel 3D B-Spline surface reconstruction technique for 3D metamorphosis for animation and entertainment. The approach allows one-to-one mapping between the object space and a parameter space, and therefore automatic correspondence between a pair of reconstructed objects. B-Spline-based shape representation also has the advantages of: 1) easy shape editing, 2) level of detail control, and 3) compact storage.

1 Introduction

3D metamorphosis is a smooth transformation from a source 3D object to a target 3D object. Unlike 2D image morphing, 3D metamorphosis is independent of viewing or lighting parameters. Therefore, it is a powerful technique for the entertainment industry. The primary task in 3D metamorphosis is to automatically establish surface correspondence between the source and target objects. By mapping each point on the source object to a point on the target object, a smooth transition can be generated by interpolating the source to the target object.

A common approach to establishing correspondence between two objects is to generate a common connectivity, which is generally accomplished by decomposing each object into several patches, embedding the patches into a 2D parametric domain, and finally merging the corresponding embeddings to form a common mesh [4,6 8,9]. This approach has three drawbacks. First, since objects are represented as dense polygon meshes, the large data set is difficult to manipulate. Second, the difference in size, scale, and topology between the source and target objects pose more problems. To handle this situation, most morphing techniques involve the use of a sparse set of user selected feature pairs to guide decomposition and to establish an initial coarse correspondence, upon which a dense correspondence can be based [7]. However, manually marking correspondence feature points on 3D objects is a difficult and tedious task. Third, the common connectivity generated is object-dependent. If either the source or target object is altered, the whole process of establishing correspondences must be repeated. Therefore, the motivation of our research is to automatically construct a compact 3D representation using B-Splines and establish a dense correspondence between objects automatically, without any user intervention. Although there has been considerable work on fitting B-Spline surfaces to 3D point clouds, the object correspondence is seldom addressed.

Z. Pan et al. (Eds.): Edutainment 2008, LNCS 5093, pp. 781–788, 2008.

In the works described in [2,5], complex surfaces are often reconstructed using a network of surface patches. Due to the uncertainty in the automatic division of surface patches, it is difficult to establish correspondences between objects. Although single B-Spline patch fitting [3] can avoid the problem of division uncertainty and eliminate the nontrivial consideration on surface continuity, previous approaches are limited to gird data representing simple topological relationships, e.g. a deformed quadrilateral region or a deformed cylinder. Our reconstruction method is not limited to grid data, and doe not demand reconstructed patches to be square shaped, so a complex object can be reconstructed on a common parameter space using much fewer patches, compared to previous approach. Detailed algorithms are presented in next section. Section 3 shows the potential applications, and conclusion is given in Section 4.

2 B-Spline Modelling

2.1 Methodology

A B-Spline curve of degree k is a weighted sum of a set of weighted control points $c_i, 0 \le i \le n$,or $C = \{c_0, c_1, \cdots, c_n\}$,

$$f(t) = \sum_{i=0}^{n} B_i^k(t) C_i \tag{1}$$

The weight $B_i^k(t)$ is a polynomial function defined over a knot vector $U = \{u_0, u_1, \cdots, u_n\}$, and is recursively calculated as follows:

$$\begin{aligned} B_i^0(t) &= \begin{cases} 1 & u_i \le t \le u_{i+1} \\ 0 & otherwise \end{cases} \\ B_i^k(t) &= \frac{t - u_i}{u_{i+k} - u_i} B_i^{k-1}(t) + \frac{u_{i+k+1} - t}{u_{i+k+1} - u_{i+1}} B_{i+1}^{k-1}(t) \end{aligned} \tag{2}$$

Similarly, a B-Spline surface is defined over a pair of knot vector $U = \{u_0, u_1, \cdots, u_n\}$ and $V = \{v_0, v_1, \cdots v_m\}$ by:

$$\Gamma(s,t) = \sum_{i=0}^{m} \sum_{j=0}^{n} N_i^k(s) B_j^r(t) c_{i,j} \tag{3}$$

Therefore, the task of B-Spline surface fitting is to find a set of control points, which defines a surface Γ giving the best approximation to a given dataset D . In other words, each point of the given data set D will be approximated by a point on the reconstructed B-Spline, as follows:

$$\Gamma(s,t) = D(s,t) = \sum \sum NBc = \sum N(\sum Bc) \tag{4}$$

Thus, the reconstruction of a B-Spline surface, i.e. the computation of a set of control points C , can actually be seen as $M \times N$ B-Spline curve-fitting processes.

However, for non-grid datasets, each curve-fitting process will be defined on a different knot vector. We propose a knot vector standardization algorithm.

Suppose F and L are two curves fitting the original non-grid dataset independently. F is defined on knot vector $X=[X_0,X_1,\cdots,X_{nx+g+1}]$ by n_x+1 control points f: $[f_1,f_2,\cdots,f_{nx}]$, L is defined on knot vector $Y=[y_0,y_1,\cdots,y_{ny+g+1}]$ by n_Y+1 control points l: $[l_1,l_2,\cdots l_{ny}]$.

$$F(x)=\sum_{i=0}^{n_x}B_{i,g}(x)f_i \tag{5}$$

$$B_{i,g}(x)=\frac{x-x_i}{x_{i+g}-x_i}B_{i,g-1}(x)+\frac{x_{i+g+1}-x}{x_{i+g+1}-x_{i+1}}B_{i+1,g-1}(x) \tag{6}$$

$$L(y)=\sum_{j=0}^{n_y}N_{j,g}(y)l_i \tag{7}$$

$$N_{j,g}(y)=\frac{y-y_j}{h_{j+g}-y_i}N_{j,g-1}(y)+\frac{y_{j+g+1}-y}{y_{j+g+1}-y_{j+1}}N_{j+1,g-1}(y) \tag{8}$$

The aim is to standardise X and Y to new knot vectors X' and Y' so that $X'=Y'=U$. Instead of simply merging all knot vectors together [10], i.e. $X'=Y'=X\cup Y=U$, resulting a large knot vector, our approach is to standardise all knot vectors to a predefined knot vector $U=\{u_0,u_1,\cdots,u_{n+g+1}\}$. The method works as follows: for each element in U and X, if $x_i\in U$, then do nothing; If $\exists k$. $(\mathrm{u_k}\in \mathrm{U})\cap(\mathrm{u_k}\notin \mathrm{X})$, insert $\mathrm{u_k}$ into X; The control points f are re-calculated as $f'=[f'_1,f'_2,\cdots f'_n]^T$ and the basis function becomes:

$$B'_{k,g}(x)=\frac{x-u_k}{u_{k+g}-u_k}B'_{k,g-1}(x)+\frac{u_{k+g+1}-x}{u_{k+g+1}-u_{k+1}}B'_{k+1,g-1}(x) \tag{9}$$

The original curve is thus:

$$f'(x)=\sum_{k=0}^{n}B'_{k,g}(x)f'_k \tag{10}$$

Similarly, control points l are re-calculated as $l'=[l'_1,l'_2,\cdots,l'_n]^T$, and the basis function is re-defined on U:

$$N'_{k,g}(y)=\frac{y-u_k}{u_{k+g}-u_k}N'_{k,g-1}(y)+\frac{u_{k+g+1}-y}{u_{k+g+1}-u_{k+1}}N'_{k+1,g-1}(y) \tag{11}$$

$$L'(y)=\sum_{k=0}^{n}N'_{k,g}(y)l'_k \tag{12}$$

From Equation 9 and 11, it can be seen that the basis functions B' and N' are identical, which can be generalised as:

$$Q_{k,g}(s) = \frac{s - u_k}{u_{k+g} - u_k} Q_{k,g-1}(s) + \frac{u_{k+g+1} - s}{u_{k+g+1} - u_{k+1}} Q_{k+1,g-1}(s) \tag{13}$$

Consequently, Equation 10 and 12 can be rewritten as:

$$F'(s) = \sum_{k=0}^{n} Q_{k,g}(s) f'_k = A(s) \cdot f' \tag{14}$$

$$L'(s) = \sum_{k=0}^{n} Q_{k,g}(s) l'_k = A(s) \cdot l' \tag{15}$$

where $A(s) = [Q_{0,g}(s), Q_{1,g}(s), \cdots, Q_{n,g}(s)]$. f' and l' are shape descriptors of curve F and L, respectively. Shape descriptors have several important properties, including:

- One-to-one mapping from the parameter domain to the object space. For each pair of parameter value (s,t), there is a unique corresponding B-Spline surface point in the object space.
- Compact representation for 3D objects. Over 90% compression rate is achieved and similar rendering result to that using original polygon representation.

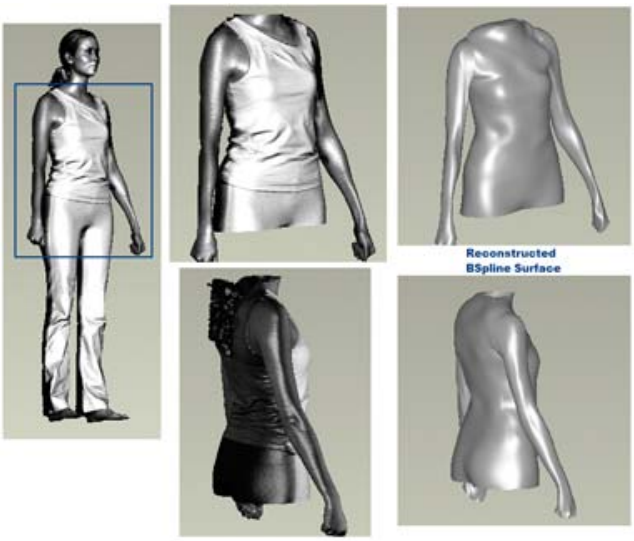
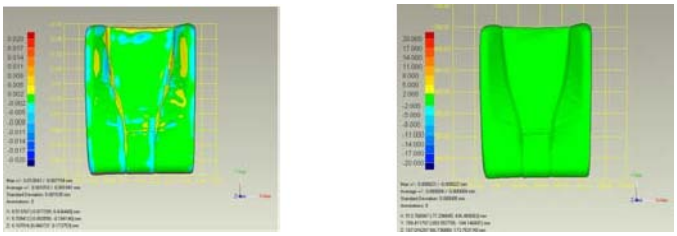


Fig. 1. Four patches B-Spline surface reconstruction example. The four patches are front and back body, left and right arms, respectively.

Figures 1, 2, 3 demonstrate some reconstructed B-Spline surfaces using our approach discussed above. For comparison, we also applied a previous B-Spline reconstruction method [5] on the same dataset in Figure 3.

Rendered single patch B-Spline surface is shown in Fig. 2 right.

Fig. 2. Textured-rendering results. Left: Polygon model (75,232 points). Right: single patch B-Spline surface model (the surface is presented by 616 shape descriptors).



(a) (b)

Fig. 3. Difference analysis. (a) Comparison between our reconstructed single B-Spline patch and the original data. The standard deviation is 0.002mm. (b) Comparison between the multiple B-Spline representation (by previous approach [5]) and the original data. The standard deviation is 0.000005mm.

Table 1. Performance comparisons between our approach and the previous B-Spline surface modelling method

Mode	Our B-Spline surface modelling	Previous B-Spline surface modelling
Vertices	75, 232	75, 232
Triangles	-----	149, 044
Modelling time	1.55 sec.	44.31 sec.

2.2 Correspondence

Similar to curve Equation 14 and 15, surface Equation 3 can be rewritten as

$$\Gamma(s,t) = A(s,t) \cdot C \tag{16}$$

Given a pair of parameter (s,t) , A is the same for every object. Thus $C = [c_{0,0}, \cdots c_{m,0}, c_{0,1} \cdots c_{m,1}, \cdots c_{m,n}]^T \in R^3$ of size of $(m+1)\times(n+1)$ defines the unique shape of a surface, i.e. C is shape descriptors. Thus, we have established a one to one mapping between the parameter domain $(s,t) \in \Omega : [0,1]\times[0,1]$ and the object space $\Gamma \in R^3$ via C . Thus the corresponding surface points between object surfaces can

then be generated. For each pair of parameters (s,t), there is a unique corresponding B-spline surface point:

$$(\mathrm{s,t}) \Rightarrow \Gamma^{\mathrm{k}}(s,t) \tag{17}$$

$$(\mathrm{s,t}) \Rightarrow \Gamma^{\mathrm{k+1}}(s,t) \tag{18}$$

Therefore, B-Spline surface points $\Gamma^{\mathrm{k}}(s,t)$ and $\Gamma^{\mathrm{k+1}}(s,t)$ are uniquely mapped, i.e.

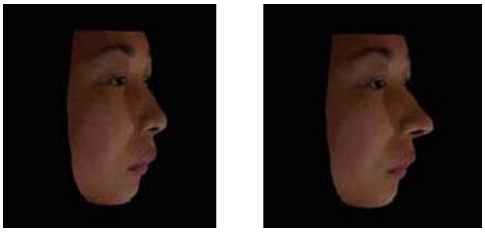
$$\Gamma^{\mathrm{k}}(s,t) \Rightarrow \Gamma^{\mathrm{k+1}}(s,t) \tag{19}$$

By sampling the parameter domain, e.g. uniform sampling, we obtain a set of corresponding B-Spline surface points on each object surface.

3 Applications

3.1 3D Morphable Model

Unlike some common techniques for data compression such as principal components analysis (PCA) describing an object as a weighted sum of principal components which often bear little resemblance to the underlying interdependent structure of biological forms [1], shape descriptors contain geometrical information about the objects. Therefore, apart from recognition purpose [11] [12], the reconstructed 3D models can be employed for freeform deformation and animation, see Fig. 4.



(a) (b)

Fig. 4. Changing facial attributes. (a) Before. (b) After.

3.2 3D Morphing

Surface-based 3D metamorphosis consists of two steps: 1) establishing a dense correspondence from surface Γ^1 to Γ^2 and 2) creating a series of intermediate objects by interpolating corresponding points between Γ^1 and Γ^2. With shape descriptors C the issue of establishing correspondences between Γ^1 and Γ^2 is simplified as calculating

$$\Gamma^1(s,t) = A(s,t) \cdot C^1 \tag{20}$$

$$\Gamma^2(s,t) = A(s,t) \cdot C^2 \tag{21}$$

where, the value of $A(s,t)$ is only related with the sampling scheme on the common parameter domain [0, 1]×[0, 1]. $\Gamma^1(s,t)$ and $\Gamma^2(s,t)$ are corresponding surface points.

Since $A(s,t)$ is the same for all objects when the same sampling scheme is adopted, it can be computed once in advance.

To create a series of intermediate objects between Γ^1 and Γ^2, we simply apply linear interpolation between corresponding points. Supposing n intermediate objects Γ_i $(1 \le i \le n)$ are required, they can be generated by

$$\Gamma_i(s,t) = \Gamma^1(s,t) + \frac{i}{n+1}(\Gamma^2(s,t) - \Gamma^1(s,t)) \tag{22}$$

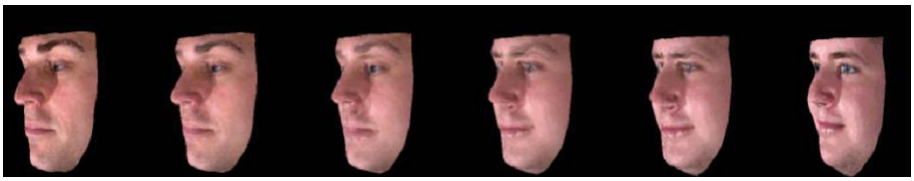


Fig. 5. Smooth 3D metamorphosis sequence

Smooth morphing from one face to another is shown in Fig. 5. Four intermediate faces are displayed between the source face (left end) and the target face (right end). Suppose the target object is changed from Γ^2 to Γ^3 and m intermediate objects are required while source object is still Γ^1, then the intermediate objects $\Gamma_j (1 \le j \le m)$ are computed as

$$\Gamma_j(s,t) = \Gamma^1(s,t) + \frac{j}{m+1}(\Gamma^3(s,t) - \Gamma^1(s,t)) \tag{23}$$

By simply changing the sampling scheme, the morphing sequence can be rendered in different resolutions. Similarly, $A(s',t')$ can be computed one time for all objects.

4 Conclusions

A novel 3D modelling technique for to 3D metamorphosis is presented. In contrast to previous works using vertices and polygons, our approach uses shape descriptors. Despite high compression rate, rendering result using shape descriptors is still similar to that using the original polygon representation. Moreover, one-to-one mapping from the object space to a common parameter space can be established and therefore correspondence between any pair of objects.

Though our approach provides a smooth and compact B-Spline representation for surfaces, it may not be the best for preserving fine geometric details of objects. To overcome this problem, a displacement map may be computed for each pixel in which an offset from a point on a surface to a point in the original raw data is recorded. Texture mapping is another economical solution to make up for the loss of fine surface

details. Since the sharp edges are rather easy to detect, they can be used to guide the object division into several patches. The single patch B-Spline modelling technique can be applied to each patch independently. In this case, although multiple B-Spline patches are needed, the amount of the patches can be significantly reduced compared to previous approaches. The nontrivial problem of enforcing G1 continuity between adjacent patches is also avoided in this case. Similar solution can be applied to objects with very complex surfaces.

References

[1] Blanz, V., Vetter, T.: A Morphable Model for the Synthesis of 3D Faces. In: Proceeding of ACM SIGGRAPH, LA, pp. 187–194 (1999)
[2] Eck, M., Hoppe, H.: Automatic Reconstruction of B-Spine Surfaces of Arbitrary Topological Type. In: Proc. 23rd Int. Conf. on Computer Graphics and Interactive Techniques SIGGRAPH 1996, pp. 325–334. ACM, New York (1996)
[3] Forsey, D., Bartels, R.: Surface Fitting with Hierarchical splines. ACM Transactions on Graphics 14(2), 134–161 (1995)
[4] Hutton, T.: Dense Surface Models of the Human Face. PhD thesis (2004)
[5] Krishnamurthy, V., Levoy, M.: Fitting Smooth Surfaces to Dense Polygon Meshes. ACM-0-89791-746-4/96/008 (1996)
[6] Kanai, T., Suzuki, H., Kimura, F.: Metamorphosis of Arbitrary Triangular Meshes. IEEE Computer Graphics and Applications, 62–75 (2000)
[7] Lee, A., Dobkin, D., Sweldens, W., Schroder, P.: Multiresolution Mesh Morphing. In: Proceedings of SIGGRAPH 1999, pp. 343–350 (1999)
[8] Lee, T., Huang, P.: Fast and Intuitive Metamorphosis of 3D Polyhedral Models Using SMCC Mesh Merging Scheme. IEEE Transactions on Visualization and Computer Graphics 9(1) (2003)
[9] Lin, C., Lee, T.: Metamorphosis of 3D Polyhedral Models Using Progressive Connectivity Transformations. IEEE Transactions on Visualization and Computer Graphics 10(6) (2004)
[10] Watt, A., Watt, M.: Advanced Animation and Rendering Techniques: Theory and Practise. Addison-Wesley, Reading (1992)
[11] Bai, L., Song, Y.: Combining Graphics and Vision for 3D Face Recognition. In: Second International Conference on Vision, Video and Graphics, Edinburgh (July 2005)
[12] Song, Y., Bai, L.: Single B-Spline Patch 3D Modelling for Facial Analysis. In: The 6th International Conference on Recent Advances in Soft Computing, Kent, Canterbury, UK (July 2006)

Author Index